David Van Hoose

The Industrial Organization of Banking

Bank Behavior, Market Structure, and Regulation

Springer

The Industrial Organization of Banking

David Van Hoose

The Industrial Organization of Banking

Bank Behavior, Market Structure, and Regulation

Springer

Professor David Van Hoose
Baylor University
One Bear Place, 98003
Waco, USA
TX 76798
David_VanHoose@baylor.edu

ISBN 978-3-642-02820-5 e-ISBN 978-3-642-02821-2
DOI 10.1007/978-3-642-02821-2
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2009937336

Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Chapter 1
Introduction

This book explores several decades of research into the *industrial organization of banking*—the study of the structure of individual banks, banking markets, and their interactions. The book has two fundamental objectives. One goal is to assist students and policymakers in climbing the field's steep learning curve as effectively as possible. The other is to provide a full survey of the field as it presently stands and thereby assist active researchers in contemplating what directions they should take the field in the future.

The book reviews recent trends in banking and surveys alternative approaches to analyzing the economics of bank decision-making. It explains different perspectives on the relationship between bank market structure and bank behavior, examines antitrust issues in banking, and assesses current understanding of the relationship between bank market structure and the stability of the banking industry. Finally, it evaluates the implications of bank capital regulation, appraises the potential interaction between market discipline and direct regulatory supervision of banks, and explores the interplay between regulation and the structure of the banking industry.

Three Fundamental Areas Within the Industrial Organization of Banking

The book focuses on three fundamental areas of study within the field of the industrial organization of banking:

1. Identifying and assessing key factors influencing decision-making by individual banks
2. Evaluating the competitive structure of banking markets and associated implications for the banking industry and society
3. Assessing the implications of proposed or actual regulations for individual banks and/or the banking industry

Each of these areas is very broad and diverse. A number of researchers contemplate issues relating to one or perhaps two of these areas but only rarely all three. It can

D. Van Hoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-642-02821-2_1, © Springer-Verlag Berlin Heidelberg 2010

prove difficult, therefore, for a student or a policymaker seeking to learn about the industrial organization of banking to locate a single source of information about the status of the field as a whole, other than individual chapters or portions of chapters in the excellent advanced banking texts by Freixas and Rochet (2008), Greenbaum and Thakor (2007), Degryse et al. (2009), and Matthews and Thompson (2005) or survey articles covering specific topic areas that are scattered across a handful of issues of academic journals and books containing collected readings.

Researchers working within any one of the three areas of the field clearly struggle to keep up to date in the other two. Perhaps as a consequence, new directions pursued within one area often fail to take into account important past or current developments within another. In theoretical research on determinants of individual bank behavior, policy-prescriptive studies sometimes overlook issues relating to interrelationships among bank-level decision-making, the market environment that the bank faces, and regulatory constraints. Naturally, ignoring such interrelationships helps in obtaining tractable results but is unlikely to yield robust predictions in relation to real-world outcomes. In addition, while practitioners of econometric work examining the structure of the banking industry recognize that they must seek to control for potential interactions among behavioral responses of individual banks, the degree of market competition, and the regulatory environment, empirical studies often abstract nonetheless from consideration of important links among bank behavior, market structure, and regulation that must govern realized outcomes within the data under consideration. Furthermore, analyses of the impacts of bank regulations commonly fail to consider how bank market structure conditions the effects of these regulations on industry performance and channels through which regulations can feed back to influence the competitive structure of the banking industry.

Objectives

This book's fundamental purpose is to assist students, researchers, and policymakers by providing a complete overview, exposition, and evaluation of the economic profession's current understanding of the interplay among bank behavior, market structure, and regulation. One key aim of this book is to assist academic professional economists and graduate students alike in developing a broad understanding of what the profession has determined about these interrelationships. Another intention is to synthesize diverse strands of the banking literature at a level appropriate for bankers and policymakers seeking to learn about the literature.

Toward these ends, the book emphasizes helping a reader to get fully up to speed on essential theories and recent empirical evidence rather than contemplating every detail of the most complex theoretical models or the most complicated econometric methods. The book thereby can serve as a springboard for those students and policymakers seeking to gain a foundational knowledge of the literature prior to engaging more advanced theories and sophisticated econometric techniques. In addition, it can function as a reference for active researchers contemplating future explorations of the interactions among bank behavior, market structure, and regulation.

The book's pedagogical approach focuses on applying basic banking models to illustrate fundamental theoretical points, concentrating on laying out key findings of empirical studies and emphasizing policy implications of both theoretical and econometric findings. Portions of the book devote attention to issues raised by the Basel II framework for banking supervision, because most bank regulators have maintained a steadfast devotion to the principles entailed in this framework even as they have recognized that events of the late 2000s undoubtedly will lead to significant revisions. The book touches at various points on developments leading up to and following the recent global financial crisis. Nevertheless, the book is not focused on these near-term issues. It has been written with a longer-term intent of providing students, policymakers, and academic researchers with a broad background on the industrial organization of banking. An extensive understanding of the field's general findings will assist readers in rethinking the appropriate competitive structure of banking markets and optimal bank regulatory configurations in light of recent experience.

Bank Behavior and the Structure of Banking Markets

Chapters 2–4 discuss the foundations of the industrial organization of banking. Chapter 2 overviews key banking concepts, including assets and liabilities, sources of income and expenses and measures of profitability, and forms of asymmetric information and risks that banks confront. The chapter also surveys recent trends in the structure of banking revealed by data from U.S. commercial banks.

Chapter 3 reviews alternative theories of bank behavior. After considering the issue of outputs versus inputs of banking institutions, the chapter examines the theory of banks as portfolio managers. It then turns to a discussion of models of banks as profit-maximizing firms incurring real resource expenses alongside the net interest revenues it earns. Considered first is a banking model that assumes the baseline case of perfect competitive behavior in bank loan and deposit markets, which is useful both for conducting static comparisons of alternative modes of competition and for explaining the important concept of portfolio separation in both static and dynamic settings. Chapter 3 then turns to the polar cases of monopoly in a bank's loan markets and monopsony in its deposit markets. Next the chapter considers standard Cournot–Nash and Bertrand–Nash models of bank behavior in oligopolistic settings with homogenous loans and deposits. The chapter concludes by discussing alternative approaches to rivalry among banks with differentiated loans and deposits in monopolistically competitive markets.

Chapter 4 applies the theories introduced in Chapter 3 to discussion of and evaluation of alternative approaches to the industrial economics of banking. Chapter 4 shows how the static imperfect-competition frameworks discussed in Chapter 3 can, along with a dominant-bank framework, be utilized to provide a foundation for the structure-conduct-performance (SCP) paradigm that many

researchers have applied in both theoretical and empirical contributions to the banking literature. It also examines empirical evidence regarding the SCP paradigm. In addition, Chapter 4 considers the interaction between bank competition and customer relationships and reviews the state of the evidence concerning this relationship. Furthermore, the chapter discusses application of the efficient-structure theory to banking and surveys the evidence regarding its applicability to real-world environments. The chapter concludes by reviewing recent work applying the theory of endogenous sunk fixed costs to the banking industry.

Bank Competition and Public Policy

Chapters 5 and 6 review fundamental policy issues associated with bank competition. Chapter 5 begins by considering rationales for bank mergers and then discusses both theoretical hypotheses and empirical evidence regarding effects of mergers on bank loans and deposits, loan and deposit rates, and social welfare. It then examines current U.S. banking antitrust policies and evaluates rationales for these policies as well as potential pitfalls in their implementation. The chapter next provides an analysis of special antitrust issues confronting card payment networks in which banks are active participants. After discussing the nature of two-sided payment networks, the chapter surveys developments in examining competition among such networks and implications for antitrust policy.

Chapter 6 focuses on the implications of market structure and competition for stability of the banking industry. Chapter 6 opens by presenting and evaluating prevailing theories of banks as issuers of demandable debt, an activity that exposes these institutions to risks of individual failures and the potential for systemic runs, thereby potentially providing a rationale for both deposit insurance and regulatory supervision. The chapter then turns to analysis of banks' special roles in intermediating informational asymmetries. In particular, the chapter explains how loan monitoring activities by banks can be incorporated into basic banking theory and reviews evidence regarding the empirical importance of bank loan monitoring activities. It concludes by discussing aspects of active governmental involvement in the banking industry intended to improve its stability prospects, including government-sponsored deposit insurance, the too-big-to-fail doctrine, and capital regulation initially established under the so-called Basel I and Basel II agreements formulated under the auspices of the Bank for International Settlements.

Assessing Bank Regulation

Chapters 7–9 examine the interplay between bank competition and regulation. Chapter 7 focuses on how industrial organization shapes the impacts of bank capital regulation formalized under the Basel I and Basel II frameworks for international

banking regulation. The fundamental message of the chapter is that alternative theories of bank behavior yield significantly different predictions regarding the effects of regulatory capital standards. Portfolio management models and incentive-based, theory-of-the-firm banking models that assume perfectly competitive banking markets produce ambiguous predictions about the safety-and-soundness impacts of capital regulation. In contrast, models emphasizing the potential for imperfect competition, particularly in bank deposit markets, tend to be more supportive of stability enhancements from capital standards, and theoretical frameworks that additionally emphasizing the potential for systemic risks and runs bank are highly supportive of stability-enhancing benefits from capital regulation. Nevertheless, taking into account bank screening and monitoring responses to capital requirements again leads to uncertain impacts of capital standards, particularly once the possibility of heterogeneous responses across banks is taken into account. The chapter concludes by considering evidence regarding the actual effects of capital standards implemented in the 1990s and 2000s and evaluating the scope for capital requirements to add to the banking industry's inherent procyclical tendencies.

Chapter 8 considers the role of market discipline in the banking industry. The chapter begins by providing a basic overview of the Basel II guidelines regarding market discipline and related conceptual issues, such as the disclosure of information, channels of market signals, and managerial responses. It reviews alternative suggestions for contributing to improved bank safety and soundness via enhanced market discipline, including proposals mandating the issuance of subordinated debts. The chapter discusses recent work aimed at integrating analysis of market discipline within a basic model of the banking firm and extends this work to analyze the relationship between bank market structure and market discipline. It then surveys the results from research assessing the extent to which markets actually discipline banks and the interaction between market discipline and supervisory discipline applied by bank regulators. The chapter closes with an evaluation of the Basel II market discipline pillar in relation to the capital standards and supervisory process pillars.

Much of the research on bank regulation presupposes that market power, asymmetric information, and/or externalities arising from systemic risks are sufficiently pervasive to justify public-interest-oriented supervision of banks. The branch of the industrial organization literature examining the economics of regulation suggests, however, that public choice rationales—interest-group desires to marshal public resources to transfer economic rents from one group to another group or to gain protection from competition for incumbent firms—also are key factors explaining regulation. Hence, Chapter 9 focuses on the interplay between bank regulation and the structure of the banking industry, recognizing that while it is true that market structure issues can be offered to rationalize regulation, it is also the case that regulation can alter the competitive structure of banking markets. The chapter explains how the economic theory of regulation can be applied to banking, thereby yielding a wide range of potential regulatory outcomes, from public-interest-oriented regulatory outcomes at one extreme to capture of bank regulators who pursue solely the interests of the regulated industry at the other extreme. After surveying research on

optimal bank closure policies, the chapter turns to consideration of the little-studied but increasingly relevant issue of competition among government regulators facing overlapping jurisdictions of clienteles that can choose which of the regulators serve as their primary supervisor. In addition, Chapter 9 reviews recent work that casts light on factors determining whether or not competition among leads to a regulatory race to the bottom in terms of the stringency and enforcement of bank supervisory standards. Furthermore, it discusses the potential for conflicts of interest facing central banks also charged with the conduct of monetary policy. In light of these public choice considerations relevant to bank regulation, the chapter evaluates the supervisory process pillar of the Basel II framework and finds it wanting. The chapter closes with an evaluation of the importance of regulatory compliance costs in banking, which it concludes constitute a significant but heretofore virtually unexplored component of endogenous sunk fixed costs in the banking industry.

Acknowledgments The contents of portions of this book have benefited from helpful comments from Jack Tatom, John Pattison, Kenneth Kopecky, Edward Kane, and several anonymous journal referees. I also thank Michael VanHoose for assistance in proofreading. I am appreciative to journal publishers for policies that permit portions of previously published articles to be incorporated into books such as this one. Chapter 5 includes parts of VanHoose (2009), Chapter 7 incorporates portions of VanHoose (2008, 2007b), and portions of Chapters 8 and 9 build on VanHoose (2007a). Finally, I am grateful to Networks Financial Institute of Indiana State University for its support of additional research that I have integrated into portions of this book. Any errors that may remain are solely my responsibility.

Chapter 2
The Banking Environment

Stocks, Flows, Information, and Risks

This chapter reviews fundamental banking concepts utilized throughout the chapters that follow. It also provides an economic assessment of recent trends in banking.

The Bank Balance Sheet

The analytical tools of industrial organization are typically applied to study the allocations of and rates of return on banks' assets and liabilities. Thus, bank balance sheets are at center stage in the industrial organization of banking.

Bank Assets

A bank asset represents a legal obligation by another party to repay principal plus any contracted interest to the bank within a specified period. Table 2.1 lists the combined assets of all domestically chartered U.S. commercial banks. There are three important asset categories listed in Table 2.1. Let's consider each in turn.

Loans

Lending is the bread-and-butter business of commercial banks, so the predominant category of assets held by commercial banks is loans. There are four important loan classifications:

- *Commercial and Industrial (C&I) Loans* Commercial and industrial loans, which Table 2.1 indicates typically account for more than 12 percent of total bank assets, are loans that banks extend to business enterprises to meet day-to-day cash needs or to finance purchases of plants and equipment. C&I loans have varying degrees of default risk and liquidity. A borrower typically must secure C&I loans with assets pledged as collateral to ensure repayment of the principal and interest on a loan. A lending bank may seize the collateral, or a portion of it, in the event of nonpayment. Although many C&I loans require collateral, it is

D. Van Hoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-642-02821-2_2, © Springer-Verlag Berlin Heidelberg 2010

Table 2.1 Assets of U.S. commercial banks

Asset category	$ Billions	%
Commercial and industrial loans	1,197.9	12.3
Consumer loans	847.4	9.0
Real estate loans	3,573.9	36.7
Interbank loans	364.6	3.6
Other loans	269.0	2.7
Total loans	6,252.8	64.3
Securities	2,017.7	20.7
Cash assets	247.1	2.5
Other assets	1,220.4	12.5
Total assets	9,738.0	100.0

(Source: Board of Governors of the Federal Reserve System, August 2008)

not uncommon for some C&I loans to extremely creditworthy borrowers to be uncollateralized.

- *Consumer Loans* Table 2.1 shows that consumer loans account for 9 percent of U.S. bank assets. In the United States, about a third of loans to consumers finance purchases of automobiles. Many individuals also obtain consumer loans for the purchase of mobile homes, durable consumer goods such as household appliances, or materials for home improvements. Banks typically issue consumer loans for purchase of autos or mobile homes through *installment credit* agreements, under which individual borrowers of consumer loans agree to repay principal and interest in equal periodic payments scheduled over a one- to five-year interval. Interest rates on these loans usually are fixed over the term of the loan, although a small portion of consumer loans have adjustable interest rates. A portion of consumer loans are extended automatically under revolving credit agreements, with the most notable example being credit card lending.
- *Real Estate Loans* These are loans that banks extend to finance purchases of real property, buildings, and fixtures (items permanently attached to real estate). From the 1980s through the late 2000s, real estate lending became a relatively more important business for commercial banks. The share of total commercial bank assets held as real estate loans rose from around 17 percent in 1985 to nearly 37 percent.
- *Interbank Loans* Finally, banks lend funds to each other directly in markets for interbank loans, such as the U.S. federal funds market in which banks borrow from and lend to each other deposits that they hold at Federal Reserve banks. Most federal funds loans have one-day maturities, though some, called term federal funds, are interbank loans with maturities exceeding one day. Banks typically extend interbank loans in large-denomination units ranging from $200,000 to well over $1 million per loan. Although large banks both lend and borrow federal funds, smaller banks predominantly are federal funds lenders.

Some loans are extended in the form of syndicated loans, which are loans pieced together by groups of banks. Typically one or two banks arrange a syndicated loan, in return for syndication-management fees. These lead banks line up a group, or syndicate, of banks that fund portions of the total amount of the loan, earning interest just as they would on any other loan they extend. Banks' shares of many syndicated loans are marketable instruments, meaning that participating banks under some circumstances can sell their shares of the loan to other banks.

Securities

As shown in Table 2.1, U.S. government securities, including Treasury bills, notes, and bonds, account for just over 20 percent of all U.S. commercial banks' assets. The other group of securities consists of state and municipal bonds, securities issues by government agencies, and mortgage-backed securities issued by firms such as the Federal National Mortgage Association.

Cash Assets

Cash assets are the most liquid bank assets that function as media of exchange. A key component of cash assets is vault cash, which is currency that commercial banks hold at their offices to meet depositors' cash requirements for withdrawals on a day-to-day basis.

The second type of cash asset is reserves held with the central bank, such as reserve deposits that U.S. banks maintain with Federal Reserve banks. Banks write checks out of or wire-transfer funds from these reserve deposit accounts when they make federal funds loans, buy repurchase agreements, or obtain securities. Funds held as reserve deposits and vault cash count toward meeting the Federal Reserve's legal reserve requirements.

Correspondent balances, or funds that banks hold on deposit with other private, correspondent banking institutions, are the third type of cash asset. The fourth is cash items in process of collection, which are checks or other cash drafts that the bank lists as deposited for immediate credit but that the bank may have to cancel if payment on the items is not received.

Trends in U.S. Bank Asset Allocations

Figure 2.1 plots the shares of bank assets allocated to cash assets, securities, and all other assets (loans and miscellaneous other assets) at various intervals since 1961. As the figure indicates, allocations to assets other than securities and cash assets—primarily loans—rose markedly into the late 1980s.

There has been a general downward trend in relative holdings of cash assets during the past several decades. Bank security holdings as a share of total assets also exhibited a slight downward trend through the latter 1980s. The percentage of assets allocated to security holdings rose thereafter, however, and remained above 20 percent of total assets to the end of the 2000s.

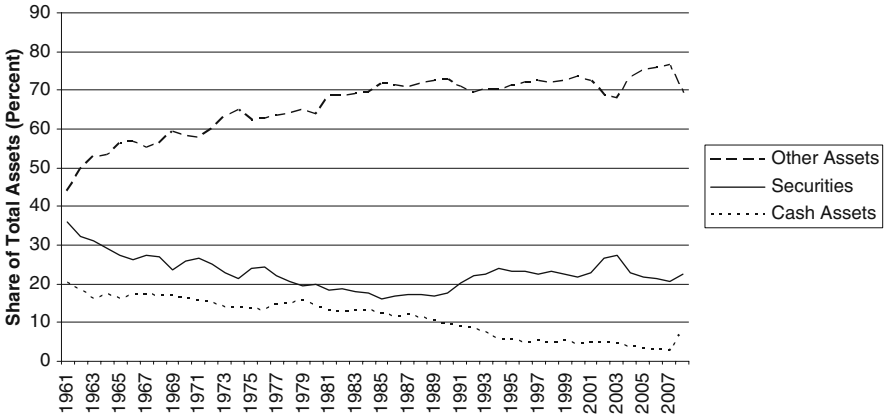

Fig. 2.1 U.S. commercial banks' asset allocations
(Source: Board of Governors of the Federal Reserve System)

Figure 2.2 displays relative allocations of U.S. bank loans to private individuals and businesses since the early 1940s. This breakdown includes agricultural loans, which constituted a significant share of bank lending in earlier years but now amount to less than 1 percent. The figure indicates that until the mid-1980s, U.S. banks had a focus on commercial and industrial loans but then diversified into real estate, interbank, consumer, and other lending. From the mid-1980s into the mid-1990s and again from the late 1990s up to the present, U.S. banks' focus shifted to real estate lending.

Do banks benefit from focusing primarily on a particular type of lending, or do they gain from maintaining a more diversified loan portfolio? Acharya et al. (2006)

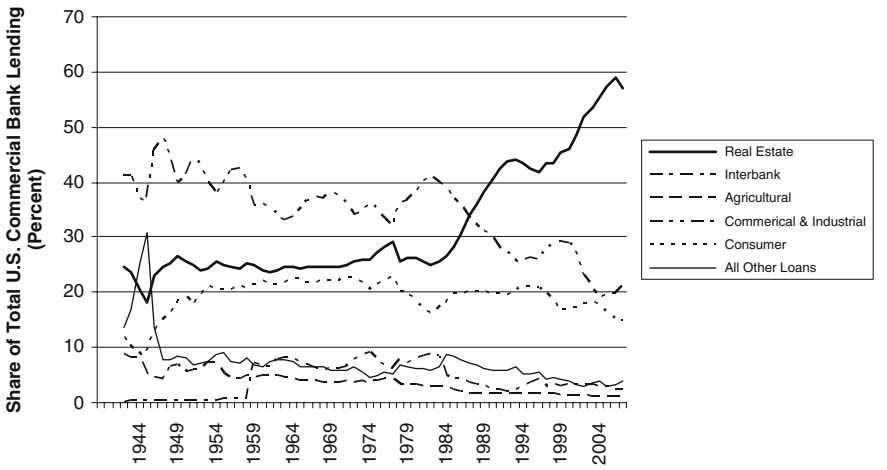

Fig. 2.2 U.S. commercial banks' loan allocations
(Source: Federal Deposit Insurance Corporation)

utilize data on returns and risk from more than 100 Italian banks during the 1990s to examine the benefits that banks derive from focus versus diversification. They conclude that diversification reduced returns of high-risk banks while increasing their lending risks. At lower-risk banks, loan diversification led to either a less efficient risk-return trade-off at best a marginal improvement in the terms of this trade-off.

Bank Liabilities and Equity Capital

A liability of a bank is the value of a legal claim on its assets. Table 2.2 lists the combined total liabilities and equity capital all U.S. banks. Let's consider each of the liability categories

Table 2.2 U.S. commercial bank liabilities and equity capital

Category	$ Billions	%
Transactions deposits	579.1	6.0
Large time deposits	1,016.4	10.4
Savings and Small Time Deposits	4,171.6	42.8
Total deposits	5,767.1	59.2
Borrowings	1,744.8	17.9
Other liabilities	1,051.4	10.8
Total liabilities	8,563.3	87.9
Equity capital	1,174.7	12.1
Total liabilities and equity capital	9,738.0	100.0

(Source: Board of Governors of the Federal Reserve System, August 2008)

Transactions Deposits

Transaction deposit accounts are accounts from which owners may draw funds via checks or debit cards. In the United States, transactions deposits include non-interest-bearing demand deposits and interest-bearing other checkable deposits. Transactions deposits account for about 6 percent of total U.S. bank liabilities and equity capital.

Large-Denomination Time Deposits

Most large-denomination time deposits, which are in denominations exceeding $100,000, are certificates of deposit (CDs) that typically fund a significant portion of banks' short-term lending operations. Large CDs pay market interest rates, and many large CDs are negotiable. Banks issue large CDs in a variety of maturities, but most large negotiable CDs have six-month terms and trade actively. All told, large CDs and other large-denomination time deposits account for just over 10 percent of bank liabilities and equity capital.

Savings Deposits and Small-Denomination Time Deposits

Included among savings deposits are passbook and statement savings accounts with no set maturities and money market deposit accounts usually held in somewhat larger denominations. Small-denomination time deposits have denominations under $100,000 and fixed maturities.

Purchased Funds and Subordinated Notes and Debentures

Key liabilities among the "borrowings" and "other liabilities" categories in Table 2.2 are purchased funds and subordinated notes and debentures. Purchased funds include interbank borrowings, central bank borrowings, Eurocurrency liabilities, and repurchase agreements.

Subordinated notes and debentures are bank debt instruments with maturities in excess of one year. Those who hold these debt instruments have subordinated claims in the event of bank failures. Thus, in the event of bankruptcy, holders of subordinated notes and debentures would receive no payments from a bank until all depositors at the bank have received the funds from their accounts.

Bank Capital

A commercial bank's equity capital is its net worth, or the amount by which its assets exceed its liabilities. As discussed in Chapters 6 and 7, bank regulators have given considerable attention to equity capital in relation to total assets. Only in recent years has the ratio of equity capital to total liabilities and equity capital—or, alternatively, total assets—risen above 10 percent.

Trends in Bank Liabilities and Equity Capital

Figure 2.3 depicts the shares of total bank liabilities and equity capital accounted for by total transactions, savings, and small and large time deposits, other liabilities, and equity capital at various dates since 1961. The figure makes clear that the general trend has been toward reduced dependence on deposit funding and a slight downward trend, until recently, in equity capital. The relative use of other liabilities, including purchased funds and subordinated notes and debentures, increased from the 1960s through the early 1980s, tended to level off in the late 1980s, and then increased considerably during the 1990s to between 20 and just over 30 percent of total liabilities and equity capital.

A key reason for the shift from deposits to purchased funds was that banks have struggled to attract sufficient deposits. Savers could earn higher yields by holding other financial instruments such as government securities, so banks borrowed from other sources to fund some of their lending operations. Raising equity funds

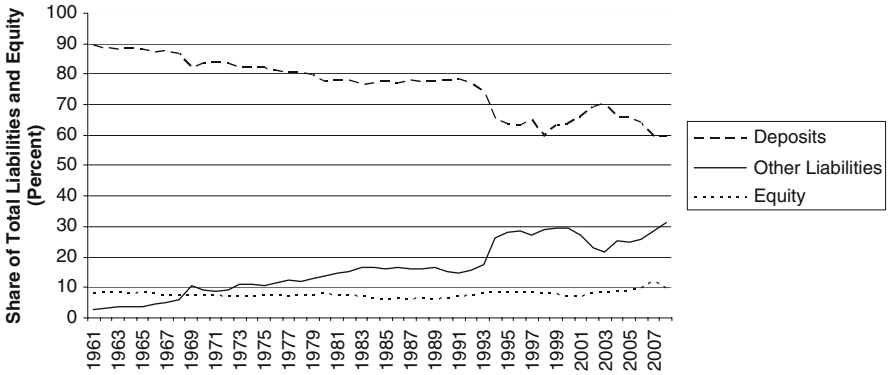

Fig. 2.3 U.S. commercial banks' liabilities and equity capital
(Source: Board of Governors of the Federal Reserve System)

in the stock market can be fairly expensive operation and can dilute the value of existing shares, so until recently banks tried to avoid issuing more stock. The main impetus for the recent change of heart concerning issuing equity capital arose from regulatory pressures that we shall discuss in detail in Chapter 7.

What difference does it make which source of funding banks utilize more heavily? Based on data from more than 1,300 banks in 101 countries between 1995 and 2007, Demirgüç-Kunt and Huizinga (2009) find that utilizing non-deposit sources of purchased funds offers risk-reducing diversification benefits at low levels of non-deposit funding. At relatively high levels of purchased funds, banks' risks of lower returns increase considerably. Mercieca et al. (2007) find no evidence of diversification benefits from heavier reliance on purchased funds at 755 small European banks between 1997 and 2003.

The Bank Income Statement

Banks measure their incomes, or revenues, as flows over time. Hence, they tabulate and report interest income in quarterly and annual income statements.

Interest Income

Figure 2.4 shows that interest income accounts for roughly two-thirds of the revenues of U.S. commercial banks. The bulk of interest income is derived from loan interest income, which accounts for just over half of total earnings of U.S. banks.

Noninterest Income

As Fig. 2.4 indicates, U.S. commercial banks earn about one-third of their revenues as noninterest income obtained from sources other than interest income, such as trading profits and customer service charges. As discussed in more detail later in this chapter, many banks sell some of the loans that they have made to other financial institutions. Such loan sales commonly include an arrangement in which the banks selling the loan continue to maintain the loan account on behalf of the purchaser. That is, they continue to manage and process payments and expenses relating to the loans even though those loans are off their books. In return for such services, banks charge fees to the loan purchasers. These loan management fees are another source of income.

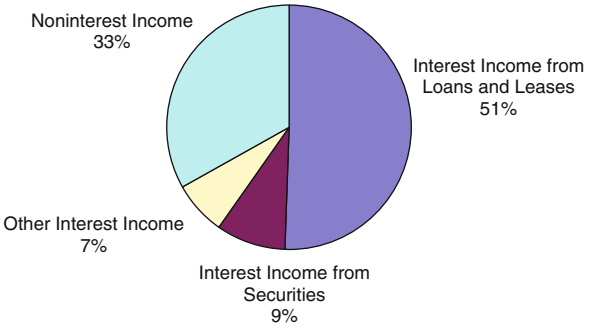

Fig. 2.4 Sources of U.S. commercial banks' revenues (Source: Federal Deposit Insurance Corporation, 2008)

Demirgüç-Kunt and Huizinga (2009) find that relying heavily on non-interest income-based activities tends to generate higher earnings volatility, a conclusion consistent with DeYoung and Roland's (2001) results derived from data from 472 U.S. commerce banks between 1988 and 1995. Furthermore, Mercieca et al. (2007) conclude that there is an inverse relationship between non-interest income and performance of across these banks, a conclusion that mirrors the results obtained by Stiroh (2004) in an analysis of the U.S. banking industry from the early 1980s through the early 2000s.

Interest Expenses

Banks apply funds raised from issuing deposits and other liabilities to acquisition of income-generating assets. To attract funds, banks must pay interest on these liabilities, and these interest expenses constitute a significant component of bank costs. As shown in Fig. 2.5, interest expenses account just over 40 percent of the total costs incurred by U.S. commercial banks.

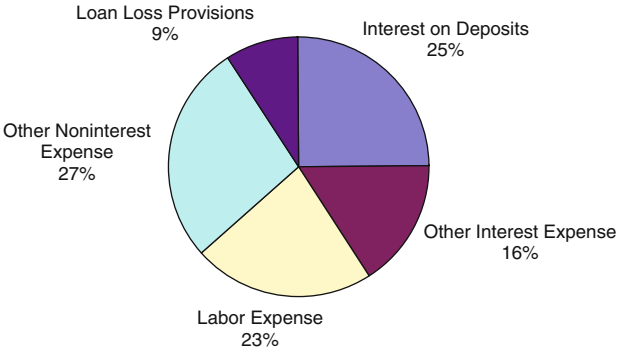

Fig. 2.5 U.S. commercial banks' expenses
(Source: Federal Deposit Insurance Corporation)

Expenses for Loan Loss Provisions

Banking is a risky business, because from time to time borrowers default on their loans. Banks earmark part of their cash assets as loan loss reserves. This portion of cash assets is held as available liquidity that banks recognize as depleted in the event that loan defaults actually occur.

Periodically, banks must add to their loan loss reserves as loan defaults cause them to decline. These additions are loan loss provisions, and they are incurred as expenses during the relevant period. Figure 2.5 shows that loan loss provisions have recently accounted for about 9 percent of expenses of U.S. commercial banks.

Real Resource Expenses

Any bank utilizes traditional factors of production—labor, capital, and land—in its operations. The bank must pay wages and salaries to its employees, purchase or lease capital goods such as bank branch buildings and computer equipment, and pay rental fees for the use of land on which its offices and branches are situated.

Figure 2.5 indicates that expenses on real resources amount to slightly over half of expenses incurred by U.S. commercial banks. Clearly, real resource expenditures are a nontrivial portion of banks' total costs.

Bank Profitability Measures

A bank's net income, or accounting profit, is the dollar amount by which its combined interest and noninterest income exceeds its total costs. For purposes of comparison of net-income performances across banks of different sizes, banking practitioners and researchers most commonly utilize three key profitability measures. One is return on assets, which is a bank's accounting profit as a percentage

of the value its assets. This performance measure is primarily an indicator of how capably a bank's management has been in transforming assets into net earnings. A second common measure relative profitability is return on equity, which is accounting profits as a percentage of the bank's equity capital. This measure of bank performance indicates the rate of return flowing to shareholders.

Return on assets and return on equity are retrospective measures of profitability that can be used to gauge relative past performances of banks. For someone aiming to gauge assess a bank's current or likely future profitability performance, a prospective profitability measure is a bank's net interest margin, which is the difference between a depository institution's interest income and interest expenses as a percentage of total assets. This profitability measure often is regarded as a useful indicator of current and future bank performance because interest income is, as shown in Fig. 2.4, such a large portion of total revenues while interest expense represents, as indicated in Fig. 2.5, a significant portion of total costs.

Figure 2.6 shows how U.S. commercial banks have performed since 1995 based on both their average return on assets and their average return on equity. All three profitability measures were remarkably stable over much of the period, until the onset of the subprime meltdown in 2007 generated sharp downturns in banks' returns on assets and equity. Net interest margin only dipped slightly prior to 2007, so it turned out to be a relatively poor prospective indicator for the late 2000s.

Berger et al. (2000) have sought to determine what factors accounted for the persistence of U.S. bank profits through the end of the 1990s. They explored a number of factors that might have accounted for this persistence, including informational opacity and banking industry competition, which are key elements of banking explored in later chapters. Their conclusion is that regional and macroeconomic shocks were consistently key determinants of profit persistence. This suggests that strong U.S. economic performance was perhaps the key factor accounting for U.S. persistent bank profitability into the 2000s, prior to the collapse of the housing market bubble in 2007 and generalized financial-markets meltdown that commenced thereafter.

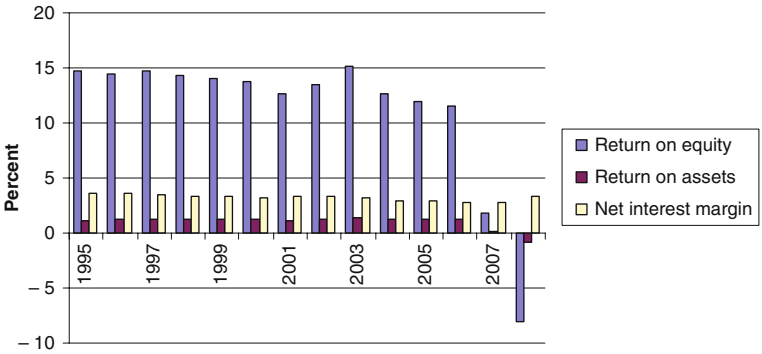

Fig. 2.6 U.S. commercial banks' average return on assets and return on equity since 1995 (Source: Federal Deposit Insurance Corporation)

As discussed by Clark et al. (2007), during the 2000s a number of banks sought to establish stronger positions in retail banking operations centered around services provided to consumers and small businesses via branch networks and the Internet. Returns on such operations tend to be more stable than returns relative to other business lines. Hirtle and Stiroh (2007) examine the U.S. banking industry between 1997 and 2004 and find that only the largest banks experience significantly reduced earnings volatility from retail banking. Those that succeeded in reducing volatility of their earnings, Hirtle and Stiroh conclude, experienced a trade-off in the form of lower returns.

Asymmetric Information and Risks in Banking

Why do so many households and firms opt to deposit funds with banks instead of lending them directly to ultimate borrowers? One key reason is the presence of asymmetric information, which arises whenever one party in a financial transaction has information not possessed by the other party

Adverse Selection

Suppose, for instance, that managers of the firm intend to utilize the proceeds of a loan to fund operations that are likely to generate a payoff more than sufficient to repay the loan. It is also conceivable, however, that the firm's managers actually have in mind allocating the funds to a project with a potentially higher payoff but also a greater likelihood of failure. From the bank's point of view, therefore, a firm seeking a loan possesses information about the intended application of desired funding that is not necessarily readily discernible. Indeed, the bank faces a danger that firms and other borrowers most interested in obtaining credit are those desiring to pursue projects with highest risks. After all, such borrowers would be gambling with the bank's funds rather than their own.

This particular asymmetric-information problem is adverse selection, or the potential that those who desire funds for undeserving projects are the most likely to seek credit. A key task that a bank confronts in lending and most of its other asset portfolio allocations is screening prospective borrowers in an effort to avoid undesired risk exposures arising from adverse selection.

Moral Hazard

Even after a bank screens prospective borrowers, identifies those deemed creditworthy, and extends credit, it faces another asymmetric-information problem. Once funds are in hand, borrowers may diverge from previously intended uses of those funds. For instance, after a bank makes a loan to a firm that had planned to apply

the funds to a relatively low-risk project, the firm's financial condition could suffer an adverse shock. In an effort to recover, the firm's managers may be tempted to re-direct the borrowed funds to a project promising a greater return but at a higher probability of failure

This possibility that a borrower may behave in a way that increases risk after a loan has been made or a debt instrument has been purchased is moral hazard. That is, after a financial transaction has taken place, a borrower can undertake actions that raise the riskiness of the financial instrument that the borrower has already issued, thereby acting "immorally" from the perspective of the lender. Thus, monitoring borrowers' actual applications of borrowed funds and on-going financial conditions is an additional key task that a bank faces as a lender.

Risks on the Balance Sheet

Because borrowers face risks of loss in operations funded by bank credit, banks confront a number of risks. Several of these risks are always present on a bank's balance sheet.

Credit Risk

A fundamental asset risk faced by a bank is credit risk, or the probability that a portion of the institution's assets—loans in particular—will decrease in value. Of course, the ultimate form of credit risk is the possibility of full default by borrowers.

Key measures of credit risk include the ratio of nonperforming loans (loans past due for at least 90 days) to total loans, the ratio of net loan charge-offs (loans declared valueless and no longer carried on the balance sheet) to total loans, and the ratios of loan loss provisions to total loans or to equity capital.

Market Risks

Because banks hold a variety of securities alongside their loans, they encounter market risks on both types of assets. One manifestation of market risk is exposure to price risk, or the potential for a sudden drop in securities prices, with a bank's degree of exposure to such risk usually measured as the ratio of the book value of assets to the estimated market value of those assets.

Another form of market risk is interest rate risk, which arises mainly through the potential for interest rates on liabilities to rise more rapidly than increases in interest rates on assets. The most common measure of a bank's exposure to interest rate risk is the ratio of interest-sensitive assets to interest-sensitive liabilities. If this ratio is significantly greater (less) than unity, then an institution is vulnerable to losses if the general level of interest rates declines (rises).

Liquidity Risk

A fundamental risk faced by banks is liquidity risk. This is the probability of having insufficient cash and borrowing capability to satisfy desired depositor withdrawals, to be able to extend loans to creditworthy borrowers, or to meet other cash requirements.

In normal times, illiquidity events are rare and isolated. When such an event takes place, an affected bank typically must borrow funds at interest rates exceeding those paid by other institutions. Liquidity risk can also arise more generally as a consequence of concerns about the stability of the banking system, which induce large numbers of depositors to seek withdrawals.

Systemic Risk

Banks assume credit, market, and liquidity risks on an individual basis. Because payment flows among banks are interdependent, however, risks confronted by individual institutions have the potential to spill over onto others. Consider, for instance, a San Francisco-based bank anticipating a wire transfer of funds from a New York bank at 1:30 eastern standard time (EST). Based on this anticipation, the San Francisco bank commits to wire funds to a bank in Chicago at 1:45 EST. The Chicago bank, in turn, agrees to wire funds to a Seattle bank at 2:00 EST, using the funds that it anticipates receiving from the San Francisco bank. Consequently, if the New York bank fails to deliver the funds promised at 1:30 EST to the San Francisco bank, the San Francisco bank may wire legal title to funds to the Chicago bank at 1:45 EST that are not really in its possession. In addition, if the New York bank discovers that some event has occurred that will keep it from sending the funds at all that afternoon, then a full chain of payment transmittals may take place even though there are insufficient funds to cover the payments.

In this example, the risk of an inability by the New York bank to settle its transaction with the San Francisco bank is a liquidity or credit risk for the latter bank. For the Chicago and Seattle banks, however, this situation constitutes systemic risk. This is a risk that some banks, such as the Chicago and Seattle banks in our example, may not be able to honor financial commitments because of payment settlement breakdowns in otherwise unrelated transactions. For these payment intermediaries, systemic risk is a negative externality, or an adverse spillover effect stemming from transactions in which they were not participants. Issues arising from spillover effects relating to systemic risks in banking are contemplated in Chapter 6.

Risks Off of Bank Balance Sheets

Banks engage in a number of activities that yield income and entail expenses and risks yet do not directly influence their balance sheets. In particular, banks extend loan commitments, securitize loans, and trade derivative securities.

Loan Commitments

A loan commitment is a promise by a bank to extend credit up to some prespecified limit under a contracted interest rate and within a given interval. As Fig. 2.7 shows, U.S. banks' commitment lending grew significantly in the 1980s before dropping off during the early 1990s and then rising once more, to roughly 80 percent of total loans.

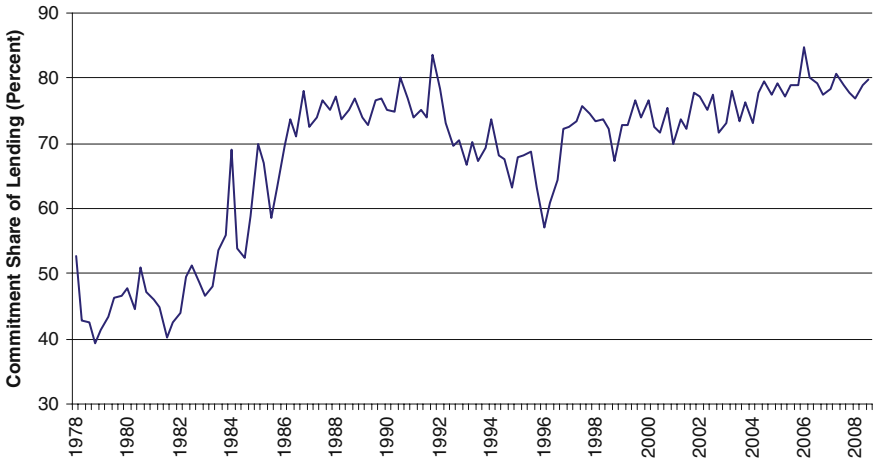

Fig. 2.7 Growth in the share of commitment lending in the United States (Source: Board of Governors of the Federal Reserve System)

In a typical loan commitment arrangement, a bank and a borrower negotiate terms of the commitment, which specify a line of credit that the bank will make available to the borrower, what the loan interest rate will be or how it will be determined, and a fee that the borrower must pay for any unused portion of the line of credit. Such an arrangement yields benefits for both the borrower and the bank. The borrower has a guarantee of credit at a given interest rate whenever desired during the specified period. The bank receives interest income on the portion of the credit line that the borrower draws upon, and the bank receives noninterest fee income on the unused portion.

There are a variety of forms of loan commitments. Under a fixed-rate loan commitment, the interest rate on any credit drawn down by the borrower is set at a predetermined level. In contrast, a floating-rate loan commitment ties the loan rate to another market interest rate, such as the prime loan rate or London Interbank Offer Rate—an average of overnight interest rates paid by major world banks. Most loan commitments of either type are revolving credit commitments. Others are confirmed credit lines, which normally provide for a bank to extend a fixed amount of credit upon demand within some short-term interval.

Securitization

Whereas a loan commitment obligates a bank to bring a loan onto its balance sheet upon a customer's request, securitization permits a bank to remove loans from a balance sheet. Securitization entails pooling loans with similar risk characteristics and selling this loan pool in the form of a negotiable financial instrument.

The first widespread use of securitization in the United States began in the early 1970s when savings institutions began selling mortgage loans to the General National Mortgage Association (GNMA), which financed its purchases of these loans by selling mortgage-backed securities to investors interested in receiving shares of the returns derived from the underlying mortgage-loan pools. Although GNMA acquired ownership of the mortgages, in return for fee income banking institutions serviced the mortgages by processing loan payments dealing with delinquency problems, and the like.

In recent decades, banks have commonly securitized many other types of loans, such as credit card loans, auto loans, and commercial and industrial loans. Securitization has enabled banks to earn fee income for originating, servicing, and insuring loans while selling them to others. Bank asset-backed securities primarily exist in two forms. One is a pass-through security, which passes interest and principal payments that a bank receives from borrowers through to investors on a proportionate basis. The second main type of bank-issued asset-backed security is a pay-through security, for which a bank initially holds the interest and principal payments for an underlying pool of loans and then reallocates them into two or more separate sets of securities that have different payment and maturity structures.

Banks can benefit from securitization because it enables them to shift credit and market risks of a portion of their lending to other parties. Securitization also generates a stable source of fee income. In addition, securitization helps to provide up-to-date information on market values of securitized loans. This can provide an indication of the market values of similar loans that a bank has chosen not securitize, which assists in the practice of valuing a bank's assets at current market values—that is, "marking to market," or market value accounting, rather than historical value accounting that would involve carrying the initial value of an asset in a bank's books until it is repaid. Of course, marking securitized assets to market can present significant downside risks, as banks discovered in the late 2000s when the market values of mortgage-backed securities plummeted when steeply declining house prices reduced the values of underlying mortgage loans.

Derivative Securities

Securitization addresses a portion of a bank's credit and market risks by moving part of its loan portfolio off its balance sheet. For a number of banks, trading derivatives also has proved to be a significant source of revenues. In one recent quarter in 2008,

for instance, revenues of $4.3 billion from trading interest rate, foreign exchange, commodities, and equity derivatives enabled U.S. commercial banks to report more than $1.6 billion in net revenues even after experiencing credit derivatives losses of $2.7 billion. By the end of 2008, U.S. banks held a notional amount of derivatives totaling more than $190 trillion, of which about $150 trillion of derivatives exposure was comprised of interest rate contracts.

Trends in U.S. Banking Industry Structure

At the heart of the study of the industrial organization of banking is evaluating effects of industry structure on banks' balance-sheet choices; on rates of return on bank assets, such as loans, and on bank liabilities, such as deposits; and on bank profits and risks. As discussed in Chapters 4 and 5, changes in banking structures have enabled researchers to explore these effects in considerable detail.

Recent Patterns in U.S. Banking Structure

Figure 2.8 shows that since the mid-1980s, the number of U.S. commercial banks has dropped by about 50 percent. This decline in the absolute number of banks has coincided with a significant change in the size distribution of the banking industry. Consider one end of this distribution, the smallest banks—often referred to in the industry as "community banks"—that each have less than $100 million in total assets. In the mid-1980s, these small banks together accounted for close to 10 percent of the combined assets of all commercial banks. Today, fewer than 40 percent

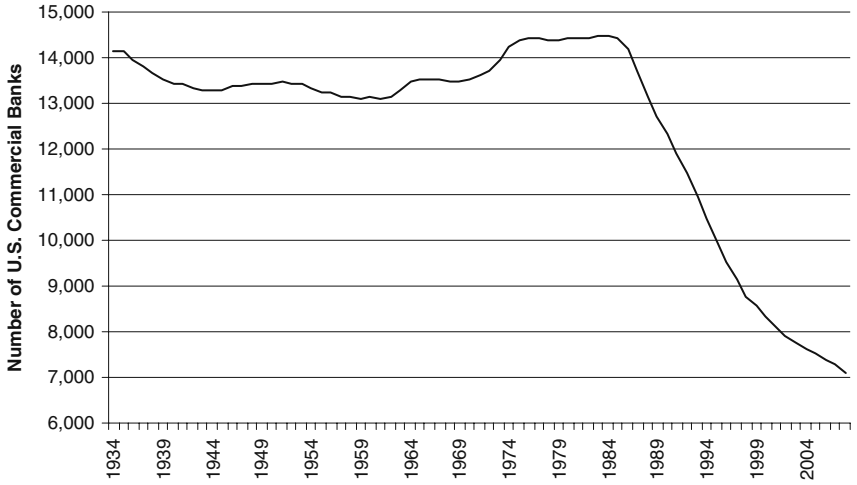

Fig. 2.8 The number of U.S. commercial banks since 1934
(Source: Federal Deposit Insurance Corporation)

of banks have total assets below $100 million, and their combined assets make up less than 2 percent of the consolidated assets of the industry.

Panel (a) of Fig. 2.9 shows that the number of bank branches has risen considerably even as the number of banks has declined. Thus, as displayed in panel (b), the average number of branches per bank has increased, further reflecting the reduction in one-branch community banks.

Mergers, Acquisitions, and Concentration

There is much more to the size-distribution story, however, than the significant drop in the relative importance of small banks. Panel (a) of Fig. 2.10 shows the annual

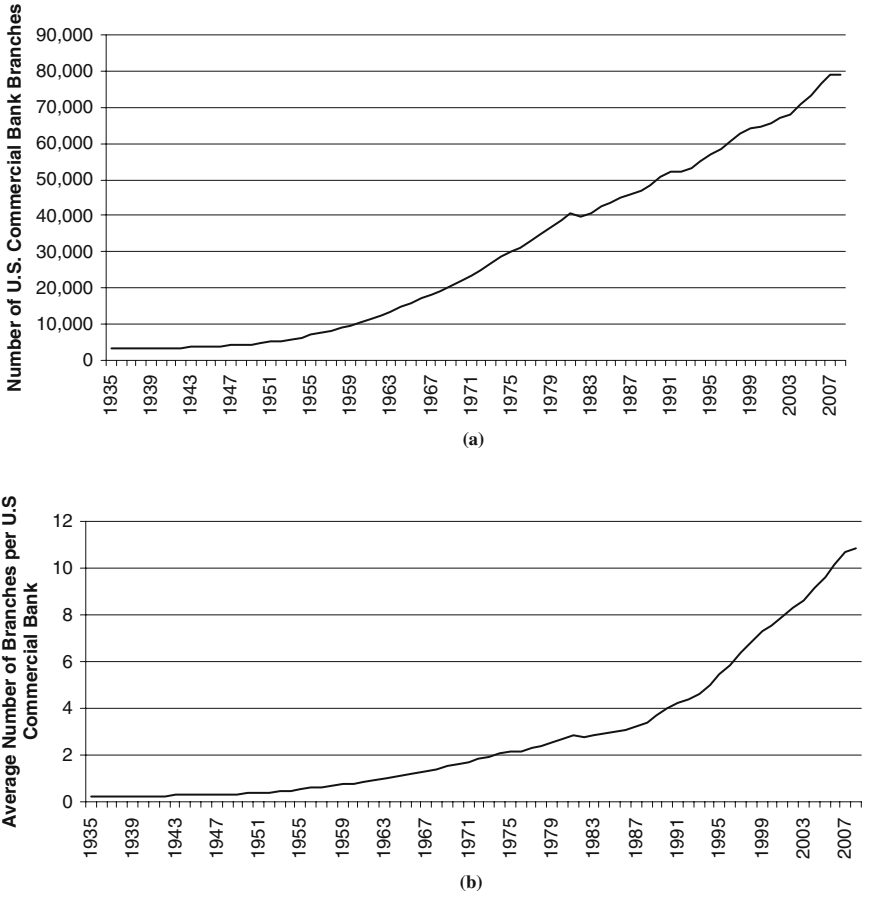

Fig. 2.9 The number of bank branches at U.S. commercial banks and the average number of branches per bank since 1934
(Source: Federal Deposit Insurance Corporation)

number of bank mergers and the total values of as sets acquired in these mergers since 1990. Each year from 1990 through 2001, mergers and acquisitions redistributed more than 2 percent of aggregate bank assets in the United States. Clearly, even though much of this merger-and-acquisition activity has involved larger banks gobbling up smaller institutions, a considerable portion also involved combinations of larger banks. Panel (b) of Fig. 2.10 shows that scores of banks and hundreds of billions of dollars of assets have been merged or acquired.

A natural consequence of this bank merger-and-acquisition wave has been an increase in aggregate industry concentration. As shown in Fig. 2.11, the percentages of deposits held at the largest U.S. banks has increased steadily since 1990. This reflects a trend, documented by Janicki and Prescott (2006) and by Jones and Critchfield (2008), of a shift in the size distribution of U.S. banks toward larger banking organizations.

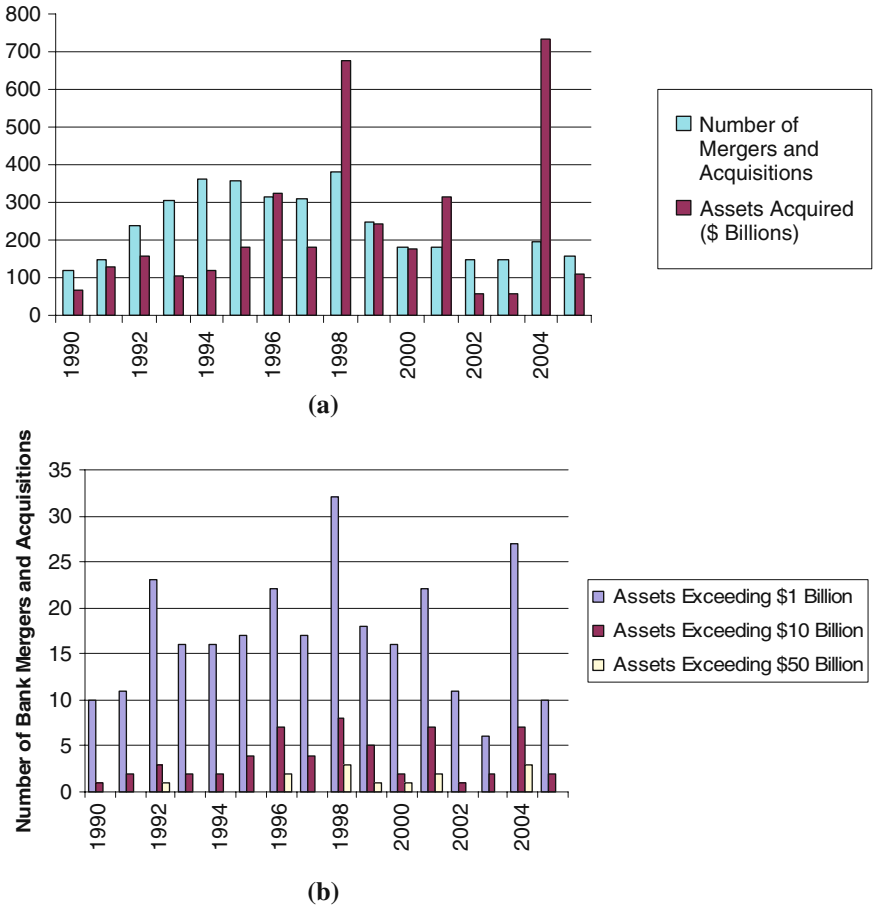

Fig. 2.10 U.S. bank mergers and acquisitions and assets acquired (Source: Pilloff (2009))

How do economists seek to take into account the relatively more concentrated nature of banking markets in efforts to understand the behavior of individual banks? What are implications of the trend toward larger banking institutions and greater market concentration for social welfare? Are more concentrated and potentially less competitive banking markets less prone to higher risks and decreased likelihood of insolvency, or do recent crisis events suggest that a less concentrated, more competitive banking industry would also be more stable? These are key questions explored in the following chapters.

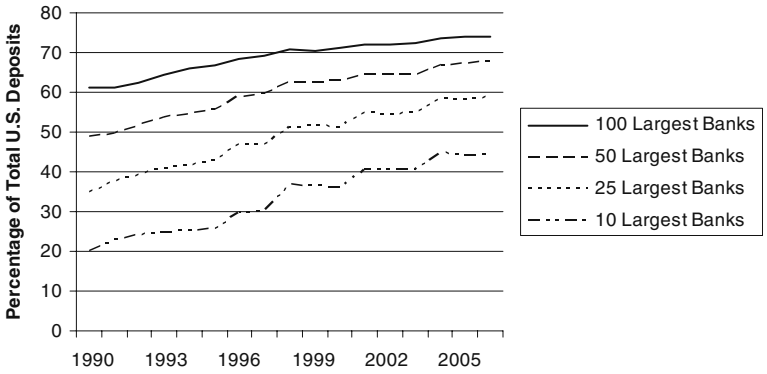

Fig. 2.11 Nationwide deposit concentration in the U.S. banking industry (Source: Pilloff (2009))

Summary: The Banking Environment

- Bank assets include loans, securities, and cash assets. In the United States, recent years have witnessed a general increase in loan assets in proportion to total assets and a significant rise in real estate lending as a share of total lending.
- Bank liabilities include transaction, savings, and time deposits and borrowings in the form of purchased funds and subordinated notes, with shares of the latter rising over time in proportion to total U.S. bank liabilities and equity capital at the expense of deposits' shares. During past decades, bank equity capital initially declined steadily relative to total assets until a recent upturn took place in response in capital-focused regulation.
- Interest income accounts for about half of the revenues of U.S. banks, although noninterest income's share of revenues has trended upward in recent years. Interest expense and loan loss reserve accruals account for about half of U.S. banks' costs. Labor expenses contribute to nearly one-fourth of U.S. banks' costs.
- Two key measures of bank profitability, return on assets and return on equity, were very stable at U.S. banks until the late 2000s, when both measures turned negative even as another profitability indicator, net interest margin, remained stable.

- Bank balance sheets are exposed to credit, market, liquidity, and systemic risks that are influenced by asymmetric information problems arising from adverse-selection and moral-hazard sources. Banks also experience off-balance-sheet sources of risk arising from commitment lending, securitization, and derivatives trading.
- Since the 1990s, scores of banks and hundreds of billions of dollars of assets have been involved in mergers and acquisitions. As a result, during the past 25 years the number of U.S. banks has dropped and the concentration of deposits among larger banks has risen even as the number of bank branches has steadily increased.

Chapter 3
Alternative Perspectives on Bank Behavior

"[T]he production process of the financial firm. . .is a multistage production process involving intermediate outputs, where loanable funds, borrowed from depositors and serviced by the firm with the use of capital, labor, and material inputs, are used in the production of earning assets."

—Sealey and Lindley (1977)

"[B]anks transform the credit portfolio demanded by borrowers into a deposit portfolio desired by lenders."

—Dewatripont and Tirole (1993)

Identifying the Outputs and Inputs of a Bank

What is a bank, exactly? All observers agree that a bank is unambiguously one among several types of financial intermediary. Such an intermediary is an institution that acts as a middleman in channeling funds from savers to entrepreneurs or other businesspeople who make capital investments or to individuals or families who purchase durable goods or tangible assets such as houses or condominiums. Savers who lend funds to financial intermediaries such as banks otherwise could have chosen to engage in direct finance by lending their funds to businesses or households without utilizing the intermediaries' services. Instead, customers of banks opt to engage in indirect finance by lending their funds to banks and other financial intermediaries in exchange for promised flows of returns on those funds. Banks and other intermediaries aim to profit from revenues derived from lending net of costs they incur by engaging in financial intermediation.

Beyond this point of agreement, researchers begin to diverge in their views on how best to define a bank. To understand why, let's consider the issues of identifying what banks produce and characterizing the markets in which they operate.

What Banks Do: Alternative Perspectives on Bank Production

The quotes above provide some indication of the difficulties involved in developing a concrete definition of a bank—and hence a single, commonly accepted theory of

D. Van Hoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-642-02821-2_3, © Springer-Verlag Berlin Heidelberg 2010

bank behavior. Dewatripont and Tirole (1993) represent a perspective that focuses on banks primarily as financial institutions that convert an asset portfolio into a set of financial instruments, namely deposits and other bank debts that surplus households and firms desire to hold in their own asset portfolios. Viewed from this perspective, banks specialize in providing a variety of financial services to savers, including the following:

- Writing and enforcing debt contracts that match savers preferring highly liquid assets with firms desiring to finance capital investment via long-term credits
- Reducing transaction costs associated with asset-liability transformation via the provision of payment services that save counterparties from incurring costs to verify their mutual solvency
- Engaging in delegated screening and monitoring to determine whether prospective borrowers are creditworthy and whether actual borrowers are directing funds to worthwhile projects
- Providing information- and risk-management services for savers

The quoted sentence from Sealey and Lindley (1977) suggests another view, in which the outputs of a bank are considered to be its earning assets, while labor and capital are physical inputs and deposits are financial inputs. According to Sealey and Lindley, customer services associated with deposits, such as payment services, represent partial payment for the use of the loanable funds provided by depositors.

As discussed by Colwell and Davis (1992) and Mlima and Hjalmarsson (2002), these two perspectives fit into two general approaches to measuring what banks produce. Under one approach, which Berger and Humphrey (1997) term the production approach and which was first utilized by Benston (1965), banks primarily specialize in producing services for holders of loan and deposit accounts. Consequently, the production approach, which receives support from the Dewatripont–Tirole discussion, recommends that appropriate measures of a bank's output should focus on numbers of various financial-service transactions performed per unit of time.

In contrast, the intermediation approach proposes that banks are primarily engaged in the process of intermediating funds between savers and borrowers. Accordingly, the intermediation approach suggests that stock values of bank assets and/or liabilities are appropriate bank output measures. Sealey and Lindley offer one particular version of the intermediation approach. They argue that only bank assets such as loans to individuals and businesses should be viewed as outputs, whereas deposit liabilities constitute inputs into an intermediation-based bank production process.

Assessing the Economic Outputs and Inputs of Banks

Berger and Humphrey (1997) suggest that the production and intermediation approaches to defining bank output can be reconciled on empirical grounds.

They note that detailed data on transaction flows are typically proprietary and unavailable to researchers. In their view, the assumption that transaction flows are proportional to the stock values of bank asset and liability accounts essentially renders both perspectives equivalent for purposes of empirical analysis of limited data.

Nevertheless, even acceptance of this conclusion leaves open the exact specification of banks' economic inputs—factors of production that cease to possess their original forms—and economic outputs—the end results of the production process by which inputs are transformed into new entities. From an empirical standpoint, three commonly used methods of identifying outputs and inputs stand out. The first is the asset method, which proposes that bank assets are output, that deposits, purchased funds, and other liabilities are financial inputs, and that real resources such as labor and capital constitute real inputs. This is the method first adopted by Alhadeff (1954) and utilized since by a number of researchers, and it accords with the theoretical arguments provided by Sealey and Lindley (1977).

The second is the value-added method, according to which a bank's outputs are identified as "banking functions which are associated with a substantial labor or physical capital expenditure to produce a (noninterest) flow of banking services" (Berger and Humphrey, 1991, pp. 125–126). This method of identifying outputs typically suggests that most key types of loans, such as commercial and industrial loans, installment loans, and real estate loans, are bank outputs. In addition, the value-added method usually identifies transactions deposits and retail savings and time deposits as outputs as well. Under this method, labor, physical capital, and purchased funds typically are classified as bank inputs.

The third is the user-cost method employed by Hancock (1985b, 1991), in which "the user cost of a financial good is defined as the net effective cost of holding one unit of services per time period" (Hancock, 1991, p. 27), which is equal to the cost of holding the asset during a current period minus the asset's discounted net revenue in the following period. Hancock classifies bank balance-sheet items with negative user costs—including all categories of loans and transactions deposits—as outputs and items with positive user costs—savings and time deposits and purchased funds—as inputs along with labor, raw materials, and physical capital.

Thus, there is a consensus in the literature that loans are unambiguously economic outputs of banks. Other candidate outputs include transactions deposit accounts and retail savings and time deposit accounts. Treating such accounts as separate "outputs" raises fundamental conceptual problems, however. Positive net values added or negative net user costs for such accounts unavoidably mix a bank's expenses on deposit funds as inputs purchased by banks and a bank's receipts from charges applied to service flows to depositors.

Consequently, henceforth assets and service flows will be regarded as the relevant outputs of banks. Deposit funds and various purchased funds will be viewed as inputs into the asset production process, and labor and capital resources will be treated as inputs into both the production of assets and the provision of service flows. Thus, the asset method of classification will be emphasized, while acknowledging the strength of Berger and Humphrey's contention that for empirical work data

limitations sometimes argue for treating certain bank deposit categories—which of course through the balance-sheet constraint will be highly correlated with banks' assets—as "output measures."

Banks as Portfolio Managers

Much of the earlier banking literature focuses on the banks as managers of portfolios of assets to which available deposit funds may be allocated. Let's begin, therefore, by examining the essential elements of this perspective on bank behavior.

The Basic Bank Portfolio-Management Model

Typical portfolio-management models of the banking firm (see, for example, Hester and Pierce, 1975) presume that a bank's owners are risk averse. In most models, owners possess a utility function characterized, at least approximately, by the first and second moments of final wealth. Thus, the owners' utility function is strictly quasi-concave and is defined over the mean, E, and standard deviation, σ, of the return on the owners' capital investment, expressed per unit of equity capital.

Consistent with standard financial allocation models, a common assumption in bank portfolio-management models is that all banks are price takers in all markets in which they operate. Thus, perfect competition prevails in all markets. Returns on assets traded in these markets are assumed to be governed by a joint—usually normal—probability distribution known by both buyers and sellers of the assets.

Following Blair and Heggestad (1978), if no risk-free asset is available to banks in light of the influence of interest rate variations on all asset returns, then banks face an efficient frontier such as EF in Fig. 3.1. This is an envelope, which may be formally derived (see, for instance, Kim and Santomero, 1988) as the solution to the problem of minimizing the variance of the overall return on a bank's diversified portfolio opportunities of mean-variance combinations attainable with returns on the set of assets to which banks may allocate deposit funds.

Under the assumption that a bank's owners derive additional utility from a higher mean return, E, but disutility from a higher standard deviation of the return, σ, indifference curves are convex. The optimal portfolio arises at a tangency of the highest attainable indifference curve, I, with the efficient frontier EF, at point P. At this point, the marginal rate of substitution between expected return and risk is equalized with the marginal rate of transformation between expected return and risk along the efficient frontier. By construction, point P corresponds to a specific allocation of various assets as per unit of bank equity on the efficient frontier. Hence, this point uniquely identifies the asset allocation that maximizes the bank owners' expected utility, which in turn reflects their underlying preferences toward expected return and risk.

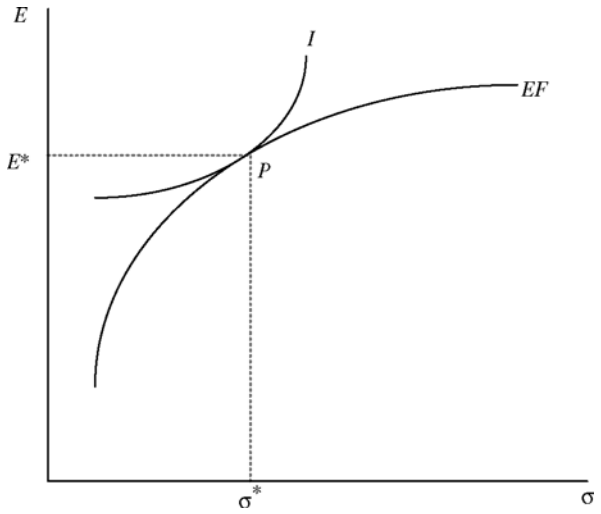

Fig. 3.1 A bank's selection of the optimal portfolio

Limitations of Portfolio Management Models

The obvious advantage of the portfolio management framework is that it represents a direct extension of basic finance theory applied to the banking firm. Naturally, the theory can be adjusted for application to special features of alternative banking environments, as considered by Szegö (1980). Indeed, as discussed in Chapter 8, the theory can reveal important implications regarding the effects of regulations of bank balance sheets, such as bank capital requirements.

Nevertheless, the assumptions underlying portfolio management models place severe restrictions on their suitability—at least, absent significant modifications—to industrial organization applications. Banks operate in a variety of markets, including markets for loans and deposits in which assumptions of standard portfolio management models—perfectly competitive price taking with symmetrically informed agents—may not even approximately apply. Indeed, in a number of policy contexts in banking, issues relating to market power and asymmetric information are of paramount importance.

Furthermore, portfolio management models of banks abstract from other important issues of concern in evaluating the industrial organization of banking. As noted in the previous chapter, more than half of the costs incurred by U.S. banks are noninterest expenses related primarily to labor and capital costs. Portfolio management models focus attention exclusively on banks' balance sheets, but banks' expenses extend beyond the balance sheet. Realistically, choices about asset allocations must be interrelated with decisions about real resource costs. Furthermore, portfolio management models typically assume a fixed scale of operations, yet in the long run a bank's scale is a choice variable, as is the distribution of sources of funds to support the selected scale.

Banks as Firms

Most modern research in the industrial organization of banking considers models of banking firms. Klein (1971) provided the first complete firm-theoretic analysis examining banks as firms utilizing inputs—funds obtained from issuing liabilities and equity capital and services of physical inputs—to produce outputs in the form of earning assets.

A Perfectly Competitive Banking Industry

Let's begin by considering a bank that operates within a perfectly competitive banking industry. Thus, this bank, which is insignificant in size relative to all markets in which it operates, issues liabilities and accumulates assets in markets in which all traded assets are homogeneous and subject to identical risks. In addition, there are no substantial barriers to entry or exit. Let's also assume that there are no informational asymmetries and that all banks are risk-neutral.

A Static Banking Model

To consider the simplest possible banking industry, let's suppose that a typical perfectly competitive bank, denoted i, has zero equity capital. (Equity could be a fixed amount carried throughout without affecting this basic analysis, at least in the absence of capital requirements, to be discussed in Chapter 7). The bank issues amounts of two liabilities with one-period maturities, deposits (D^i) and non-deposit liabilities (N^i), and uses these liabilities to fund acquisition of two single-period, interest-earning assets, loans (L^i) and government securities (S^i). Deposits potentially are subject to a reserve requirement, $R^i \geq qD^i$, where q is the required reserve ratio that may be specified by a central bank or other governmental banking authority with the power to assess a reserve requirement. Let's consider the case in which the minimum reserve requirement is binding, so that $R^i = qD^i$. Hence, the bank faces the balance-sheet constraint, $L^i + S^i = (1–q)D^i + N^i$.

In a perfectly competitive market, the bank takes as given the rates of return it pays on its liabilities (r_D and r_N) and that it earns on its assets (r_L and r_S). Thus, its interest expenses during a single period are given by $r_D D^i + r_N N^i$, and its interest earnings are $r_L L^i + r_S S^i$. Consequently, the bank's net interest margin during the period is $(r_L L^i + r_S S^i – r_D D^i – r_N N^i) / (L^i + S^i + R^i)$.

Of course, the bank must also expend real resources in raising liability funds and providing services to holders of these liabilities, and it must incur costs in screening and monitoring loans and in managing its security portfolio. Let's suppose that these costs are captured by an implicit cost function, $C^i(L^i, S^i, D^i, N^i)$, with $C^i_Z \equiv \partial C^i/\partial Z^i \geq 0$, for $Z^i = L^i$, S^i, D^i, and N^i, so that marginal costs of expense-generating activities associated with assets and liabilities are positive. In addition, let's assume that these marginal resource costs are generally increasing, so that $C^i_{ZZ} \equiv \partial^2 C^i/\partial(Z^i)^2 \geq 0$, but that resource costs are separable in individual

balance-sheet choices, so that $C^i_{ZY} \equiv \partial^2 C^i / \partial Z^i \partial Y^i = 0$, for $Y^i \neq Z^i$. Taking into account resource costs, the individual bank's profits are equal to $\pi^i = r_L L^i + r_S S^i - r_D D^i - r_N N^i - C^i(L^i, S^i, D^i, N^i)$.

Perfectly Competitive Markets for Bank Assets

Panels (a) and (b) of Fig. 3.2 provide a diagrammatic exposition of the determination of the amount of lending by a profit-maximizing bank that takes all asset and liability rates as given. In panel (a), an individual bank's marginal return on lending is the market clearing loan rate r^*_L. The bank's total marginal cost of lending, denoted MC^i_L, equals the sum of the interest rate per dollar of deposits available to lend, the marginal resource cost of nondeposit-liability funds available to lend, and the marginal resource cost generated by allocating that dollar to lending, or $MC^i_L = r_N + C^i_N + C^i_L$. This schedule is the individual bank's loan supply schedule. Summing across all banks' MC^i_L schedules yields the market supply of loan funds by the competitive banking industry, given by L^s_b in panel (b). The market clearing loan rate, r^*_L, arises at the crossing point of this schedule with the market demand schedule for loan funds across borrowers among the nonbank public (households and firms), denoted L^d_p. This loan rate, which corresponds to the individual bank's marginal revenue (MR^i_L) derived from each additional dollar of lending, equates the quantity of loans demanded and supplied, denoted L^*. The profit-maximizing quantity of lending by bank i is $L^{i,*}$, the quantity of lending at which marginal revenue equals marginal cost.

Banks are also suppliers of funds in the market for government securities. In panel (a) of Fig. 3.3, an individual bank's marginal return on government securities is the market government security rate r_S. The bank's total marginal cost of its portfolio of government securities, denoted MC^i_S, equals the sum of the interest rate

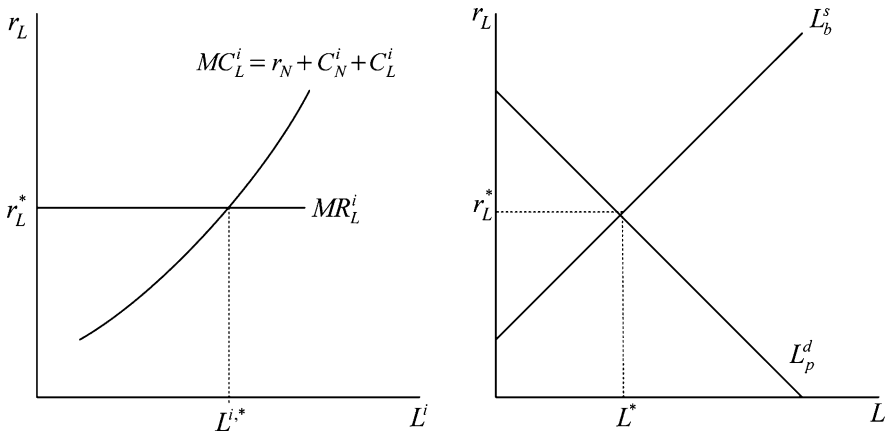

Fig. 3.2 A perfectly competitive bank loan market

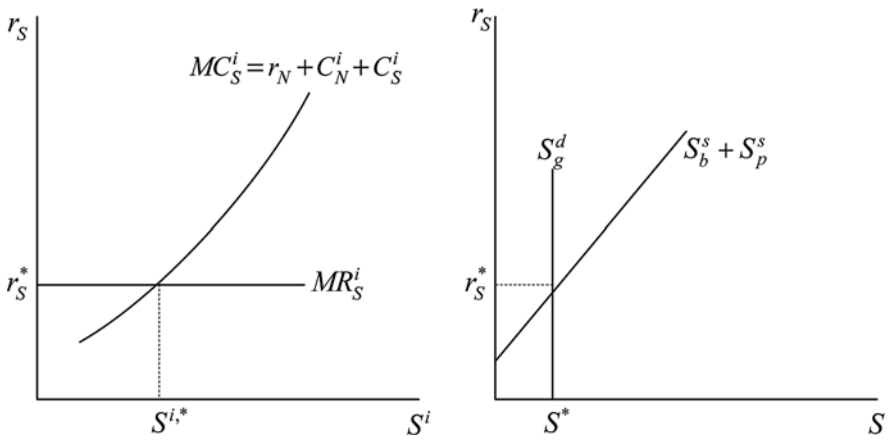

Fig. 3.3 The market for government securities

per dollar of nondeposit-liability funds available to allocate to government securities, the marginal resource cost of these funds available for such allocation, and the marginal resource cost generated by allocating an additional dollar of funds to securities, or $MC_S^i = r_N + C_N^i + C_S^i$, which is the individual bank's supply schedule of funds to the securities market. The sum over all banks' MC_S^i schedules yields the market supply of funds to the government securities market by the competitive banking industry given by S_b^s in panel (b). Adding the supply of funds to the government securities market on the part of the nonbank public, S_p^s, yields the total market supply of funds. The demand for funds by the government issuer of securities, S_g^d, is assumed for the sake of simplicity to be perfectly inelastic for simplicity in panel (b) of Fig. 3.3. Together, the combined supply of funds to the securities market bank banks and the public and the demand for funds by the government determine the market clearing rate on government securities, r_S^*. At this security rate, the profit-maximizing quantity of securities held by bank i is $S^{i,*}$. The total equilibrium quantity of securities held by banks and the public equals S^*, the quantity issued by the government.

Perfectly Competitive Markets for Bank Liabilities

Figure 3.4 displays a perfectly competitive market for bank deposit funds. In panel (a) of the figure, an individual bank takes the market clearing deposit rate as given. The bank's net marginal return on a dollar of deposit funds, NMR_D^i, is equal to the difference between the net return on a dollar of funds available to be held as assets, such as securities, $(1–q)\ (r_S - C_S^i)$, and the marginal resource cost of deposits, C_D^i, or $NMR_D^i = (1–q)\ (r_S - C_S^i) - C_D^i$. This is the individual bank's derived demand for deposit funds used as an input in the production of earning assets. Summing over all banks' NMR_D^i schedules yields the market demand for deposit funds, denoted D_b^d in panel (b) of Fig. 3.4. At the crossing point with the nonbank public's market supply

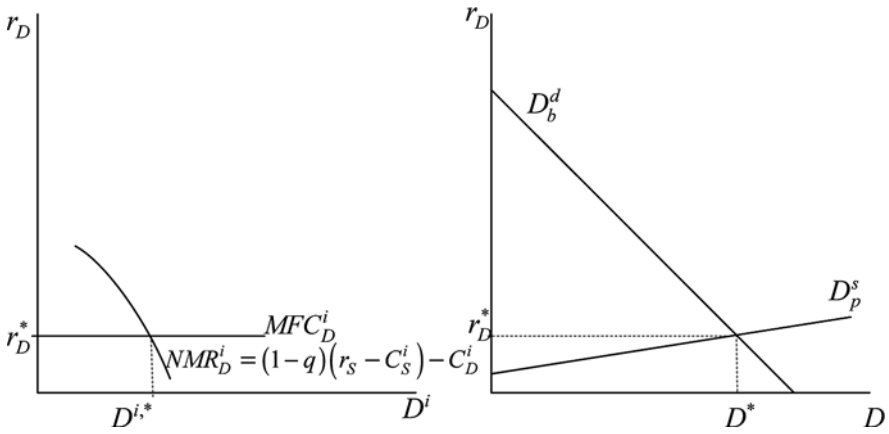

Fig. 3.4 A perfectly competitive bank deposit market

of deposit funds, D_p^s, the equilibrium total quantity of deposits, D^*, and the market deposit rate, r_D^*, are determined. The latter is the individual bank's marginal factor cost (MFC_D^i) of each deposit dollar. Equalization of NMR_D^i and MFC_D^i yields the bank's profit-maximizing quantity of deposits, $D^{i,*}$.

Finally, an individual bank takes the rate it must pay on funds raised by issuing nondeposit liabilities as given. As shown in panel (a) of Fig. 3.5, its net marginal return on a dollar of such funds, NMR_N^i, equals the difference between the net return on a dollar of these nondeposit funds held as assets, $r_S - C_S^i$, and the marginal resource cost of nondeposit liability funds, C_N^i.This difference, $NMR_N^i = r_S - C_S^i - C_N^i$, is the derived demand for funds raised from issuing nondeposit liabilities to fund the production of earning assets. Summing over all banks' NMR_N^i schedules yields the market demand for nondeposit-liability funds, denoted N_b^d in panel (b) of Fig. 3.5. The quantity of nondeposit-liability funds, N^*, and market rate of return on these funds, r_N^*, are determined the crossing point with the nonbank public's market supply of these funds, N_p^s. The market return is the individual bank's marginal factor cost, MFC_N^i, which the bank equalizes with NMR_N^i to determine the profit-maximizing quantity of nondeposit-liability funds, $N^{i,*}$, to issue.

Evaluating Properties of a Static Perfectly Competitive Banking System

Note that the marginal conditions implied by panels (a) in Figs. 3.2–3.5 imply that for any bank in this model perfectly competitive banking system, the following condition must hold true:

$$r_L - C_L^i = r_S - C_S^i = (1-q)^{-1}\left(r_D + C_D^i\right) = r_N + C_N^i$$

Thus, the optimally configured bank balance sheet is one in which net marginal returns on funds allocated to the bank's loan and security assets, $r_L - C_L^i$ and $r_S - C_S^i$,

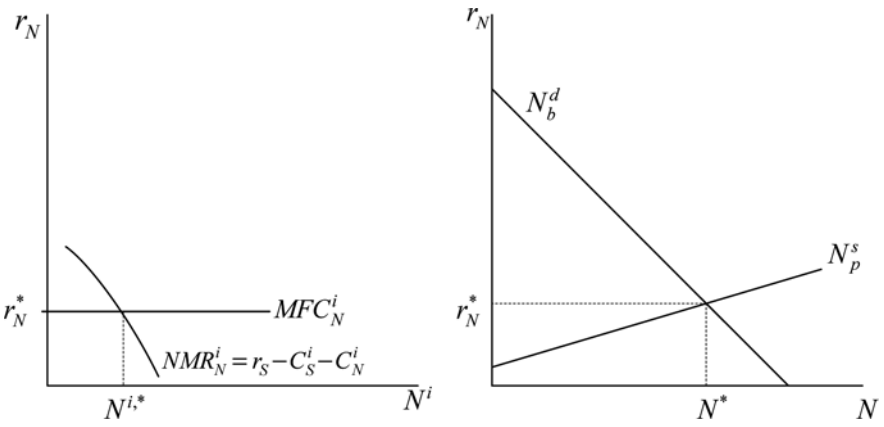

Fig. 3.5 The market for nondeposit-liability funds

are equalized with total marginal interest and non-interest expenses on deposit and nondeposit liabilities, $(1-q)^{-1}$ $(r_D + C_D^i)$ and $r_N + C_N^i$. Of course, C_L^i, C_S^i, C_D^i, and C_N^i all depend on the balance-sheet choices of the bank, so at an optimum the bank's choices insure that the values of these marginal resource costs satisfy this condition. Equalization of net funds returns on the asset side of the bank's balance sheet with net marginal funding costs on the liability side, of course, parallels the standard "price equals marginal cost" result in standard competitive theory, implying balance-sheet choices consistent with allocative efficiency in a perfectly competitive banking system.

Note that the above condition implies that, in general, a bank's asset and liability decisions are interdependent. Suppose that the above equality does not hold. As a specific example, consider the case in which the equation above initially holds, but then the loan rate rises as a result of an increase in market loan demand in panel (b) of Fig. 3.2. In comparison with the initial profit-maximizing outcome, the bank's lending is now too low. The bank responds by raising its loans, which pushes up $C^{\mathrm{i}}{}_{\mathrm{L}}$ somewhat; nevertheless, $r_{\mathrm{L}} - C_{\mathrm{L}}{}^{\mathrm{i}}$ remains higher than before. The bank therefore cuts back on security holdings, which reduces $r_{\mathrm{S}} - C_{\mathrm{s}}{}^{\mathrm{i}}$ toward equality with the new, higher value of $r_{\mathrm{L}} - C_{\mathrm{L}}{}^{\mathrm{i}}$. In addition, the bank raises more deposit and nondeposit-liability funds, which pushes up $C_{\mathrm{D}}{}^{\mathrm{i}}$ and $C_{\mathrm{N}}{}^{\mathrm{i}}$ and hence raises $(1-q)^{-1}$ $(r_{\mathrm{D}} + C_{\mathrm{D}}{}^{\mathrm{i}})$ and $r_{\mathrm{N}} + C_{\mathrm{N}}{}^{\mathrm{i}}$ to equality with this higher value of $r_{\mathrm{L}} - C_{\mathrm{L}}{}^{\mathrm{i}}$. Similar reasoning with respect to variations in other variables taken as given by an individual perfectly competitive bank—in this model, other market clearing interest rates and the required reserve ratio—likewise imply adjustments across the bank's entire balance sheet. Thus, in this basic banking model, a bank's asset and liability decisions must be interdependent. Such interdependence in the face of a higher market loan rate would play out in Figs. 3.3–3.5 by accompanying shifts in the market supply of bank funds to the government securities market and in the market demands for deposit and nondeposit liability funds, resulting in higher interest rates on securities, deposits,

and nondeposit liabilities. Market interest rates thereby would move together, as we typically observe following disturbances such as shocks to the public's loan demand.

The early literature on bank behavior (see, for instance, Pringle, 1973, Miller, 1975, Sealey, 1977, Baltensperger, 1980, and Swank, 1996) devoted considerable attention to the issue of asset-liability interdependence, in light of the conclusion in earlier work by Klein (1971) indicating that banks' asset allocations could be examined separately from liability funding decisions. Sealey (1977) noted that this *portfolio separation* result followed from Klein's assumption that a bank's asset demand and liability cost conditions are functions of the *ratios* of individual balance-sheet amounts to total assets and liabilities. Baltensperger (1980) argued that the inclusion of real resource costs also overturns portfolio separation. Both authors' points are taken into account in the above model, which is expressed in terms of *levels* of loans, securities, deposits, and nondeposit liabilities and which includes a role for resource costs.

Indeed, Baltensperger (1980) is correct that once the bank's problem is expressed in choices of levels, accounting for real resource costs in a non-trivial manner—as noted by Santomero (1984, p. 588), by incorporating the bank's "output mix as a critical determinant of operating expense"—is a crucial factor in determining whether portfolio separation holds. To see this, suppose that the nondeposit liability in the competitive banking model is interbank loans, such as federal funds, which the bank can either borrow or lend, so that N^i can take on either a positive or negative value. In addition, suppose that $C_N^i \rightarrow 0$, so that the marginal resource cost of interbank borrowing or lending is insignificant. In this case, the profit-maximizing condition across the bank's balance sheet segments into the following *set* of conditions:

$$r_L - C_L^i = r_N; r_S - C_S^i = r_N; (1-q)^{-1}(r_D + C_D^i) = r_N.$$

Consider now the example of the effect of an increase in the loan rate. Again, the bank expands lending, and C_L^i increases in value. Now, however, the bank's balance sheet constraint is satisfied by a reduction in net federal funds lending (or an increase in federal funds borrowing) without any requirement for the bank to adjust its securities holdings or its issuance of deposit liabilities. Portfolio separation holds true.

It turns out that this portfolio-separation result is one special case of the general set of conditions that Sealey (1985) identified for banks and other depository intermediaries: (1) shareholder unanimity regarding portfolio decisions, (2) separability of a bank's resource cost function, and (3) bank access to a market for funds with equal ex post borrowing and lending rates. In the model above, in which banks and hence their owners are assumed to be risk-neutral, shareholder unanimity is not an issue. Furthermore, the resource cost function is assumed to be separable. Thus, letting $C_N^i \rightarrow 0$ and allowing the bank either to borrow or lend in an interbank market does indeed yield portfolio separation for our model banking industry.

Whether or not bank asset and liability decisions are interdependent—that is, whether or not the portfolio separation result obtains in real-world banking

systems—is a non-trivial issue. From a theoretical point of view, properties of detailed banking models utilized for study of industrial organization, regulation, and other policy issues often are much easier to analyze if a researcher assumes that portfolio separation holds. Nevertheless, if asset and liability decisions are assumed to be independent from one another, spillover effects from one market to another are ruled out. For instance, altering the model to allow for imperfect competition in, say, the loan or deposit market (as considered shortly) has impacts only in that particular market; there are no spillover impacts on other bank decisions. Alternatively, in extensions of the perfectly competitive banking model to monetary policy analysis (see, for instance, Benavie and Froyen, 1982), the channels through which central bank actions such as open market purchases in the securities market influence banks' choices become much more limited under portfolio separation than in a setting in which banks asset and liability choices are completely interdependent.

Fundamental Dynamics in a Perfectly Competitive Banking Model and Implications for Portfolio Separation

Static models will prove useful throughout this book for contemplating a number of key issues in the industrial organization of banking. Nevertheless, real-world banks undeniably operate in dynamic settings. In many policy contexts, therefore, dynamic issues can prove very important. As discussed by Perloff et al. (2007), there are two basic sources of dynamics in industrial organization models. One is strategic dynamics, which arise in imperfectly competitive settings in which two or more firms' output and/or input decisions are interdependent. Another source of dynamics, which is applicable to a perfectly competitive market, is interactions among fundamental determinants of product demands and/or production processes over time. In banking, therefore, fundamental dynamics might arise as a consequence of intertemporal dependencies in the nonbank public's demands for and/or supplies of funds that create dynamic patterns in market interest rates. Alternatively, a fundamental source of dynamics could arise from banks' utilization of quasi-fixed inputs or output production processes in which output adjustments occur gradually across time.

Indeed, Flannery (1982) provided evidence that funds that many retail customers place on deposit at banks are quasi-fixed inputs, so that banks face intertemporal costs of adjustment for such deposits. Furthermore, Cosimano (1987, 1988) and Cosimano and Van Huyck (1989) examine banking models in which *either* intertemporal deposit adjustment costs *or* analogous adjustment costs of lending serve as sources of fundamental dynamics in the banking industry. Following Elyasiani et al. (1995), *both* deposit *and* loan adjustment costs can be taken into account by considering a dynamic representation of the above perfectly competitive banking model. Following the latter authors, let's simplify somewhat by assuming that the bank's only assets are loans, but let's consider the time $t + j$ profit function,

$$\pi^i_{t+j} = r_{L,t+j}L^i_{t+j} - r_{D,t+j}D^i_{t+j} - r_{N,t+j}N^i_{t+j} - C(L^i_{t+j},D^i_{t+j},N^i_{t+j}),$$

where N^i_{t+j} again is interpreted as net interbank borrowing and where, as in Cosimano (1987, 1988), Cosimano and Van Huyck (1989), Humala-Acuna (2005), and Hülsewig et al. (2006), the implicit cost function is assumed to be quadratic:

$$C(L^i_{t+j},D^i_{t+j},N^i_{t+j}) = (\alpha_1/2)\left(L^i_{t+j}\right)^2 + (\alpha_2/2)\left(L^i_{t+j} - L^i_{t+j-1}\right)^2 + \delta(N_{t+j})^2 + (\theta_1/2)\left(D^i_{t+j}\right)^2 + (\theta_2/2)\left(D^i_{t+j} - D^i_{t+j-1}\right)^2.$$

According to this specific, separable implicit cost function, in which all parameters are nonnegative constants, banks face both contemporaneous costs governed by the α_1,δ and θ_1 parameters and intertemporal deposit and loan adjustment costs determined by the α_2 and θ_2 parameters.

An individual bank in this dynamic version of the perfectly competitive model maximizes the expected discounted present value of its intertemporal flow of current and future profits, $V^i_t = E_t \sum_{j=0}^{\infty} \beta^j \pi^i_{t+j}$, where E_t is the expectations operator conditioned on time t information, and β is the discount factor, which lies in the interval (0, 1), subject to the balance-sheet constraint, $L^i_{t+j} = (1-q)D^i_{t+j} + N^i_{t+j}$. Elyasiani et al. (1995) show that as long as all parameters in the cost function, including δ, are positive, the value-maximizing bank's contemporaneous level of lending depends on the amount of loans it extended both in the last period and on the lending by the bank extended two periods previously. In addition, its contemporaneous loans depend on lagged, contemporaneous, and expected future interest rates on loans, deposits, and net interbank borrowings. Likewise, the contemporaneous value-maximizing level of deposit funds demanded by the bank depends on deposits in the last and previous periods and on lagged, contemporaneous, and expected future rates on loan, deposit, and net interbank borrowings.

The one-period lags in loans and deposits result from an "own" adjustment-cost effect; for instance, when the bank determines optimal contemporaneous lending, it takes into account the cost of adjusting its loan level relative to loans in the previous period. The second-period lag results from interdependence of loan and deposit levels. For example, when the bank determines its contemporaneous loan level, it recognizes that the "own" one-period lagged adjustment in lending will also create a lagged adjustment cost in its deposits, and it optimally smooths the costs arising from its intertemporal balance-sheet adjustments by adjusting its lending over two periods as well as by adjusting its lending to interest rates that prevailed in the previous period, hence the lagged, contemporaneous, and expected future interest-rate effects on contemporaneous lending. An analogous adjustment process arises in the bank's setting of its contemporaneous level of deposits.

As in the static competitive banking model, the intertemporally adjusting banks in this dynamic setting face costs that are separable in their balance-sheet choices. Banks are also risk-neutral, so shareholder unanimity issues considered by Sealey (1985) do not arise. Furthermore, banks can either borrow or lend interbank funds, so if $\delta \to 0$, banks can borrow or lend at the same unit cost given by the interbank

rate. Paralleling the conclusion in the static model, therefore, the portfolio separation result holds. Elyasiani et al. (1995) demonstrate that the value-maximizing contemporaneous lending level with $\delta \to 0$ depends only on deposits in the previous period, because under portfolio separation only "own" adjustment costs influence the bank's contemporaneous lending. Additionally, only the contemporaneous and expected future loan and interbank interest rates affect the contemporaneous loan choice. Contemporaneous deposits analogously depend only on deposits in the prior period and on the contemporaneous and expected future deposit and interbank rates. Hence, there is no interdependence between the bank's loan and deposit decisions across time, just as in the static model there was no interdependence within a single period.

Elyasiani et al. (1995) exploit the differences in predicted responses of contemporaneous loans and deposit for $\delta > 0$ versus $\delta \to 0$ by conducting empirical tests on quarterly loan and deposit data for 76 U.S. banks from the first quarter of 1981 through the second quarter of 1991. They find that consistent with portfolio separation, contemporaneous loans and deposits responded to one-period-lagged levels and to contemporaneous "own" interest rates and the interbank rate. They find additional evidence, however, of significant responses to two-period-lagged levels of loans and deposits, consistent with interdependence of dynamic loan and deposit choices, but evidence of cross responses to interest rates and responses to lagged interest rates is more mixed. Overall, their results cast some measure of doubt on the plausibility of the portfolio-separation assumption.

DeYoung and Yom (2008) examine more recent banking data and reach a contrary conclusion. Although they do not relate their work to Baltensperger (1980), Sealey (1985), or any other theoretical work on the theory of the banking firm, they advance a parallel argument regarding portfolio separation between a bank's asset and liability choices. DeYoung and Yom suggest that banks effectively can accomplish portfolio separation via the use of interest rate swaps and other strategies aimed at reducing risks arising from maturity mismatches not taken into account in static or dynamic models in which banks hold assets and liabilities with maturities of a single period. In effect, the use of such instruments permits banks to satisfy Sealey's condition for portfolio separation under which a bank effectively can either borrow and lend at the same per-unit rate in an asset market. Based on 1990–2005 data from FDIC-insured commercial banks, DeYoung and Yom find evidence that banks that heavily utilize interest-rate swaps come closer to exhibiting portfolio separation than institutions that do not. They also find some evidence of greater independence of asset and liability decisions over time, particularly at larger banks that have more heavily adopted strategies for utilizing such instruments as risk-mitigation devices.

Imperfectly Competitive Banking Markets

Only recently have empirical studies, such as Molnár (2008) and Martín-Oliver (2009) sought to apply modern microeconometric techniques to estimation of joint bank decisions without presupposing portfolio independence. These authors allow

for rate-setting behavior by banks, however. Thus, they consider the possibility that banking markets may be imperfect.

Monopolistic and Monopsonistic Interest Rate Determination in Bank Loan and Deposit Markets

As will be discussed in Chapter 5, in most U.S. urban and rural bank deposit markets, measured concentration is at or above levels that raise merger concerns among antitrust authorities who regulate nonfinancial industries. If greater concentration is indeed associated with imperfectly competitive behavior in banking markets, what are the economic implications?

To answer this question, let's return to a static banking framework. Within that setting, let's continue to assume that the markets for government securities and nondeposit liabilities are perfectly competitive. Consider, however, a situation in which only one bank faces the nonbank public's demand for loan funds and supply of deposit funds. Under this setting with a loan-market monopoly and deposit-market monopsony, the bank's profit function is $\pi^i = r_L^i(L^i)\ L^i + r_S S^i - r_D^i(D^i)\ D^i - r_N N^i - C^i(L^i, S^i, D^i, N^i)$, where $r_L^i(L^i)$ and $r_D^i(D^i)$ are the inverse loan demand and deposit supply functions the bank faces, with $\partial r_L^i/\partial L^i < 0$ and $\partial r_D^i/\partial D^i > 0$.

In the general case of portfolio interdependence, the following condition now must be satisfied for a profit-maximizing bank:

$$\left[1 - (\varepsilon^i)^{-1}\right] r_L^i - C_L^i = r_S - C_S^i = (1-q)^{-1}\left\{\left[1 + \left(\eta^i\right)^{-1}\right] r_D^i + C_D^i\right\} = r_N + C_N^i,$$

where ε^i is the absolute value of the elasticity of loan demand with respect to the loan rate, $\left|\dfrac{(\partial L^i/L^i)}{(\partial r_L^i/r_L^i)}\right| > 1$, and η^i is the elasticity of deposit supply with respect to the deposit rate, $\dfrac{(\partial D^i/D^i)}{(\partial r_D^i/r_D^i)} > 1$.

Figure 3.6 depicts the determination of this bank's profit-maximizing loan and deposit rates, quantity of loans, and quantity of deposit funds issued. In panel (a), the bank is a monopolist in the loan market. It takes the funds issued in the perfectly competitive nondeposit-liabilities market as given and faces the market loan demand schedule and the associated marginal revenue curve. The profit-maximizing bank lends to the point at which marginal revenue, $[1 - (\varepsilon^i)^{-1}]r_L^i$, equals total marginal cost, $r_N + C_N^i + C_L^i$. Thus, at an optimum, the monopoly loan rate can be viewed as a markup over the marginal cost of raising funds and incurring associated resource costs, of $r_L^{i,M} = [1 - (\varepsilon^i)^{-1}]^{-1}](r_N + C_N^i + C_L^i)$.

In panel (b) of Fig. 3.6, the bank is a monopsonist in the market for deposit funds used as an input in asset transformation. The bank issues deposits to the point at which the net marginal return from raising a dollar of deposits to allocate to securities holdings, $(1 - q)(r_S - C_S^i) - C_D^i$, equals the marginal factor cost of rate of return on allocating the available portion of the marginal deposit

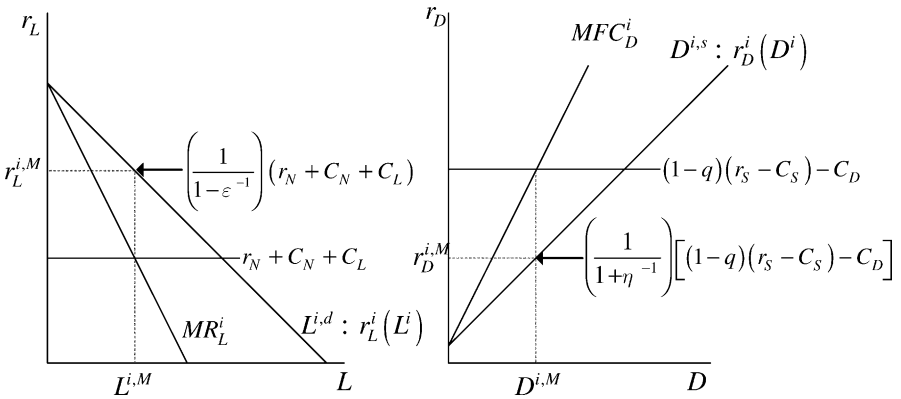

Fig. 3.6 Monopolistic loan and deposit markets

dollar to securities, given by the quantity $[1 + (\eta^i)^{-1}]r_D^i$. Consequently, the profit-maximizing deposit rate selected by the monopsonist is a markdown from the reserve-requirement-adjusted security rate and related marginal real resource costs, or $r_D^{i,M} = [1 + (\eta^i)^{-1}]^{-1}[(1-q)(r_S - C_S^i) - C_D^i]$.

Of course, banks may lend or issue deposits in more than one market. Extending the model to include multiple imperfectly competitive loan and deposit markets would imply, as shown for instance by VanHoose (1985) and Hannan (1991), that for any two markets for loans denoted loan 1 and loan 2, the ratio of loan rates charged by the monopoly bank will be $\frac{r_{L,1}^i}{r_{L,2}^i} = \frac{[1-(\varepsilon_2^i)^{-1}](r_N + C_N^i + C_{L,1}^i)}{[1-(\varepsilon_1^i)^{-1}](r_N + C_N^i + C_{L,2}^i)}$. The loan rate that a monopoly bank charges in loan market 1, therefore, is more likely to be higher than the loan rate charged in loan market 2 if the loan demand elasticity is lower and the marginal resource cost of lending is higher in loan market 1 than in loan market 2.

In any two markets for deposits denoted deposit 1 and deposit 2, the ratio of deposit rates offered by the monopsony bank will be $\frac{r_{D,1}^i}{r_{D,2}^i} = \frac{[1+(\eta_2)^{-1}][(1-q_1)(r_S - C_S^i) - C_{D,1}^i)]}{[1+(\eta_1)^{-1}][(1-q_2)(r_S - C_S^i) - C_{D,2}^i)]}$. Thus, the deposit rate paid by a monopsony bank in deposit market 1 is more likely to be lower than the deposit rate offered in deposit market 2 if the deposit supply elasticity is lower, the required reserve ratio is higher, and the marginal resource cost of raising funds is higher in deposit market 1 than in deposit market 2.

Social Losses Due to Imperfect Competition in Banking

As in any industry, the presence of monopoly and monopsony power creates social losses. Consider Fig. 3.7, which is drawn under the simplifying assumption that $C_N^i \to 0$, so that $r_S - C_S^i = r_N$, and hence the portfolio separation condition is

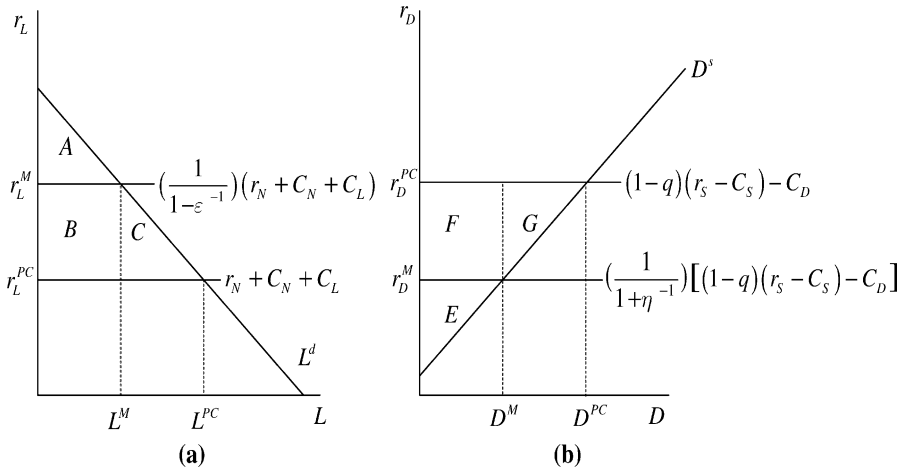

Fig. 3.7 Social losses with constant marginal resource costs

satisfied, consistent with empirical evidence provided by Adams et al. (2002) based on U.S. banking data between 1987 and 1996 indicating imperfect competition but separability of bank loan and deposits markets. Additionally, let's assume for simplicity that C_S^i, C_L^i, and C_D^i have constant values (that is, $C_{SS}^i, C_{LL}^i, C_{DD}^i \rightarrow 0$) that are the same across banks. Under these assumptions, as shown in panel (a), the total marginal cost faced by a bank—or by any set of banks, including a large set of perfectly competitive banks—in the loan market is constant. If this market were perfectly competitive, the total marginal cost schedule, $r_N + C_N + C_L$, would be the average cost schedule as well and additionally would correspond to the sum of such schedules across all banks, or the market loan supply schedule. The equilibrium loan rate would be r_L^{PC}, and the equilibrium quantity of lending would be L^{PC}. The sum of the areas A, B, and C would represent consumer surplus accruing to borrowers.

A monopoly bank, however, extends an amount of loans equal to L^M in panel (a) and charges a loan interest rate equal to r_L^M. Thus, the rectangular area B, the maximized profits earned by the monopoly bank, represent a transfer of a portion of consumer surplus to the bank. In addition, because the monopoly bank restrains lending, the triangle C represents a portion of consumer surplus that no longer obtains in the market, or a deadweight loss to society. Consumer surplus declines to the triangular area A.

Under the imposed assumptions, $(1-q)(r_S - C_S) - C_D$, the net marginal return on deposits in panel (b) of Fig. 3.7, is constant. If this market were perfectly competitive, this net marginal return schedule would be the average return schedule as well and additionally would correspond to the sum of such schedules across all banks, or the market deposit demand schedule. The equilibrium deposit rate would be r_D^{PC}, and the equilibrium quantity of lending would be D^{PC}. The sum of the areas E, F, and G would represent producer surplus yielded to depositors.

A monopsony bank, however, offers only the rate r_D^M and restrains the quantity of deposit funds it raises to D^M. It thereby receives a transfer of producer surplus away from depositors that is equal to the rectangular area F in panel (b) of Fig. 3.7. Because fewer deposits are issued, there is also a deadweight loss of producer surplus equal to the triangular area G. The triangular area F is the amount of producer surplus that remains.

Thus, a bank loan-market monopolist restricts the quantity of loans to the profit-maximizing level, which yields a loan rate above the perfectly competitive level, a transfer of consumer surplus to the bank, and a deadweight loss of consumer surplus. A bank deposit-market monopsonist limits issuance of deposits, which yields a deposit rate below the perfectly competitive level, a transfer of producer surplus to the bank, and a deadweight loss of producer surplus.

Alternative Modes of Behavior between Perfect Competition and Monopoly and Monopsony

Realistically, pure monopoly or monopsony is a rare occurrence in any modern setting. A bank typically faces at least a few market rivals, although banking markets may not be perfectly competitive, either. Consequently, theories of intermediate market structures are often usefully applied to the banking industry.

Oligopoly and Oligopsony in Banking Markets

One simple approach to examining banking markets between the extremes of perfect competition and monopoly is a Cournot–Nash framework—that is, a setting assuming quantity rivalry among banks offering homogeneous products—based on VanHoose (1985), which in turn builds on Dasgupta and Stiglitz (1981). Suppose that an individual bank i is one of m banks competing in the market for loans and $n$rivals in bank deposit markets. From this bank's perspective, total market loans are $L = L^i + \hat{L}^i = \sum_j^m L^j$, where L^i is the amount of loans extended by bank i and $\hat{L}^i$ is the exogenous (from this bank's perspective) quantity of loans extended by all other banks. Likewise, the total amount of deposits is $D = D^i + \hat{D}^i = \sum_k^n D^k$, where D^i is the amount of loans extended by bank i and $\hat{D}^i$ is the exogenous (again, from this bank's point of view) quantity of loans extended by all other banks.

Finally, suppose that loans and deposits of all banks in the respective markets are viewed by bank borrowers and depositors as homogeneous, that the inverse loan market demand function is given by $r_L = \delta r_S^{-\lambda} L^{-(1/\varepsilon)}$, and that the inverse deposit market supply function is given by $r_L = \beta r_S^{\alpha} D^{(1/\eta)}$, where $\alpha, \beta, \delta, \varepsilon, \eta$, and λ are nonnegative parameters and where ε and η are the absolute market elasticities of loan demand and market elasticity of loan supply. In this setting, bank i's profit-maximizing condition becomes

$$\left(1-\frac{L^i}{L}\varepsilon^{-1}\right)r_L^i - C_L^i = r_S - C_S^i = (1-q)^{-1}\left(1+\frac{D^i}{D}\eta^{-1}\right)r_D^i + C_D^i = r_N + C_N^i.$$

.

Finally, let's assume again that C_S^i,C_L^i, and C_D^i are constants and identical across banks. For the m banks in the loan market, this condition can be arranged and summed across all banks in that market to yield the market loan rate:

$$r_L = \Omega(r_N - C_N - C_L), \text{ where } \Omega = \frac{m}{m-\varepsilon^{-1}},$$

and for the n banks in the deposit market, rearranging and summing across all banks in the deposit market implies the market deposit rate:

$$r_D = \theta[(1-q)(r_S - C_S) - C_D], \text{ where } \theta = \frac{n}{n+\eta^{-1}}.$$

On one hand, these two expressions yield the monopoly loan rate r_L^M displayed in panel (a) of Fig. 3.7 if $m = 1$ and the monopsony deposit rate r_D^M in panel (b) of that figure if $n = 1$. On the other hand, the expressions yield the perfectly competitive loan rate r_L^{PC} in panel (a) if $m \to \infty$ and the perfectly competitive deposit rate r_D^{PC} in panel (b) if $n \to \infty$. For intermediate, finite values of m in the loan market, the equilibrium loan rate lies between the extremes of pure monopoly and perfect competition, as does the equilibrium quantity of lending. Thus, the transfer of consumer surplus to banks and the deadweight loss are smaller than in the pure monopoly case. Likewise, if the value of n is greater than unity and finite, the market deposit rate is in a range between the extreme situations of pure monopsony and perfect competition, and the transfer of producer surplus to banks and the deadweight loss are smaller than in the pure monopsony situation. Naturally, these implications of the Cournot–Nash banking framework suggest that society is better served by encouraging "large" numbers of rivals in bank loan and deposit markets, and hence low degrees of concentration.

In the context of this model, what happens if entry into the banking industry is costless, so that banks enter to the market to the point at which economic profits are extinguished? VanHoose (1988a) has examined this question in the context of a version of this model in which markets for all bank assets, including loans, are all perfectly competitive. On the one hand, as just noted, as n increases, total industry deposits increase, which yields the deposit-market equilibrium schedule labeled *DE* in panel (a) of Fig. 3.8. On the other hand, as industry deposits rise along the market deposit supply schedule, industry revenues initially increase, which implies that an increase in the number of banks, n, can occur while maintaining zero economic profits at individual banks. Eventually, however, further increases in deposits must ultimately result in declining revenues, so the number of banks must eventually decline to keep economic profits at zero. Thus, the zero-profit schedule labeled *ZP* in panel (a) is convex. As long as the elasticity of deposit supply with respect to the deposit rate is sufficiently low, the point of intersection of the *DE* and *ZP*

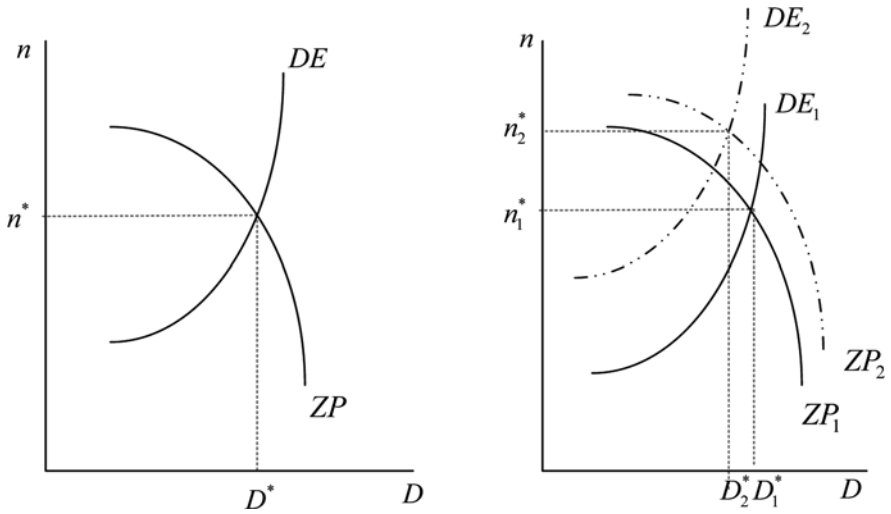

Fig. 3.8 Zero-profit equilibrium in the Cournot–Nash banking model

schedules depicted in panel (a) is a stable zero-profit equilibrium corresponding to the maintained assumption of constant marginal cost.

Panel (b) depicts the effects arising if there is an increase in the constant marginal cost of intermediation in this banking industry. Deposits across the banking industry must decline at any given deposit rate to maintain deposit-market equilibrium, so the *DE* schedule shifts leftward. At the same time, each bank responds by reducing its scale of operations, implying a rise in the number of banks consistent with a zero-profit equilibrium for a given level of industry deposits and hence an upward shift in the *ZP* schedule. The net effects are a reduction in total industry deposits and an increase in the number of rival banks operating at smaller deposit scales.

It is arguable that borrowers and savers may not always view rival banks' loans and deposits as homogeneous. There may be bank product differentiation. In an environment of differentiated loan outputs and deposit inputs, each bank faces the demand for its own individual loans and deposits, distinguishable from those of its competitors. In such an environment in an oligopolistic setting, a Bertrand–Nash framework may be applicable. Under the assumption of Bertrand–Nash behavior, banks choose loan and deposit interest rates and service fees recognizing that their interest rate and fee choices elicit responses from their rivals. For instance, in the case of a banking loan-market duopoly, suppose that bank 1 confronts the loan demand function, $L_1^d\left(r_L^1, r_L^2\right)$, with $\frac{\partial L_1^d}{\partial r_L^1} < 0; \frac{\partial L_1^d}{\partial r_L^2} > 0$, and that bank 2 analogously faces $L_2^d\left(r_L^2, r_L^1\right)$, with $\frac{\partial L_2^d}{\partial r_L^2} < 0; \frac{\partial L_2^d}{\partial r_L^1} > 0$. Each bank's profit-maximizing loan-rate reply depends positively on the other's loan-rate choice, with the magnitude of its reply dependant on the own and cross-price elasticity of demand for loans. In the limit as the banks' products become more fully

substitutable and hence homogeneous, equilibrium interest-rate replies approach perfectly competitive outcomes.

Monopolistic and Monopsonistic Competition in Banking Markets

Another approach to examining bank behavior in a setting with differentiated loans and deposits is to adopt Chamberlin (1962)—style monopolistic/monopsonistic-competition model. In this setting, there are sufficiently large numbers of competitors in banking markets that banks are individually too relatively small to affect one another's choices, although average decisions by all other banks do impact each bank's choices. An example of this type of model is the framework utilized by Startz (1983). Within the context of the present framework, let's assume that banks do not issue nondeposit liabilities, so $N = 0$, but and that they can either borrow or lend at the same per-dollar cost in an aggregate securities market in which both government and private securities trade side by side, which implies that $C_S \rightarrow 0$. This implies that portfolio separation holds. Consequently, each bank's profit-maximizing conditions for loans and deposits is tied separately to the government security rate.

Given these background assumptions, suppose that each bank faces an identical demand for loans by borrower given by

$$L^i = \frac{A_0}{m} + \frac{a}{m-1}\sum_{i \neq j}^{m}(r_L^j - r_L^i) - \frac{A_L}{m}r_L + \frac{A_S}{m}r_S,$$

where A_0, a, A_L, and A_S are nonnegative constants. Summing this expression across the m banks in the monopolistically competitive loan market yields the market loan demand schedule,

$$L = A_0 - A_L r_L + A_S r_S.$$

The latter relationship implies that the nonbank public regards loans and government securities as gross substitutes and hence reduces the quantity of loans demanded in response to a higher average market loan rate, r_L, and in response to a lower security rate. According to the former relationship, however, an individual bank i's share of lending also depends on the spread between loan rates of rivals and its own loan rate.

Analogously, the individual bank's supply of deposits from the nonbank public is, if there are n bank rivals in the monopsonistically competitive deposit market, given by

$$D^i = \frac{B_0}{n} - \frac{b}{n-1}\sum_{i \neq k}^{n}(r_D^k - r_D^i) + \frac{B_L}{n}r_D - \frac{B_S}{n}r_S,$$

with the market deposit supply schedule defined by the sum of this relationship across the n banks:

$$D = B_0 + B_D r_D - B_S r_S.$$

As in Startz (1983), a linear form for the implicit cost function might be proposed in order to compute solutions for the loan and deposit rates selected by individual banks, which given the assumption of identical banks equal the market rates on loans and deposits. A standard feature of this type of monopolistic/monopsonistic competition model is that as the number of banks competing in loan and deposit markets increases, the market loan and deposit rates become more responsive to the security rate. In the limiting cases in which $m \rightarrow \infty$, the market loan rate becomes $r_L = r_S + C_L$, which in light of the simplifying assumptions is the perfectly competitive loan rate. For $n \rightarrow \infty$. the perfectly competitive deposit rate, $r_D = (1 - q)r_S - C_D$, emerges as the equilibrium outcome.

An alternative approach to analyzing monopolistic/monopsonistic competition in the banking industry is to employ the spatial product-location models of Hotelling (1929) and Salop (1979) Examples of applications of this approach to banking include Besanko and Thakor (1992), Chiappori et al. (1995), Matutes and Vives (2000), and Cordella and Yeyati (2002). In the context of the basic framework we have been examining, the essential elements of this approach can be illustrated within a risk-neutral version of the Besanko and Thakor model. Let's suppose that there are L identical borrowers, each of which desires an amount of credit normalized to a single unit of funds. In addition, there are D identical depositors, each of which wish to save an amount also normalized at unity. There are n banks, each of which faces the same balance-sheet and cost conditions assumed in the Chamberlin-style monopolistic/monopsonistic competition model discussed above. Each bank possesses attributes differentiating it from its competitors, which may include differences in the nature of their service quality, geographic proximity, and other characteristics of importance to potential customers. These attributes are assumed to be distributed uniformly along a circle with a circumference equal to one, implying that the separation between attributes, measured by the distance between each of the banks along this unit circle, is equal $1/n$. Thus, an increase in the number of banks reduces the degree of differentiation among banks.

The L borrowers and D depositors are arranged uniformly around the unit circle. Figure 3.9 depicts the situation faced by a prospective borrower located a distance l_i^j from bank i's attribute location whose next-closest bank in terms of attributes is bank j. This borrower faces a cost equal to tl_i^j when borrowing or depositing with bank i. The borrower will choose to obtain credit from bank i instead of bank j when the total cost of borrowing from bank i, $r_L^i + tl_i^j$, where r_L^i is bank i's loan rate, is less than or equal to the total cost of borrowing from a bank j located $\frac{1}{n} - l_i^j$ units distant, given by $r_L^j + t\left(\frac{1}{n} - l_i^j\right)$.

The borrower will be willing to borrow from bank i as long as the cost of doing so is less than the cost of borrowing from bank j, or when $l_i^j \leq \dfrac{1}{2n} + \dfrac{r_L^j - r_L^i}{2t}$. Of course, the same argument applies with respect to a borrower located on the opposite

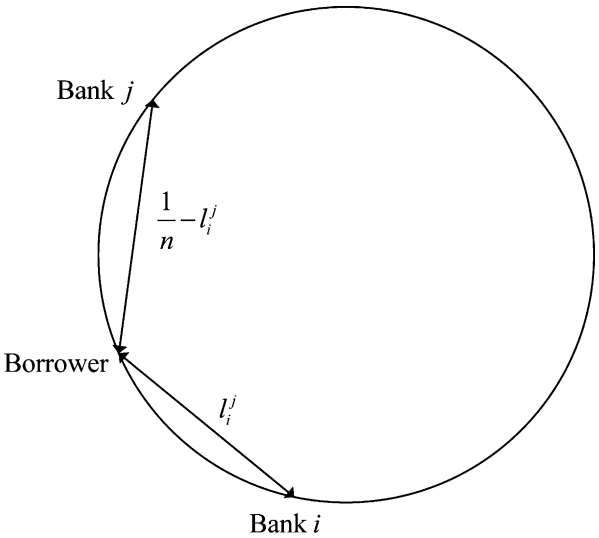

Fig. 3.9 Borrower and bank locations on a Salop circle

side of bank i whose alternative nearest bank is bank k, yielding $l_i^k \leq \frac{1}{2n} + \frac{r_L^k - r_L^i}{2t}$ for that borrower. Hence, bank i is able to extend loans to the fraction of all funds demanded by borrowers in the market, equal to L, that lies within the range along the unit circle given by $\left[l_i^j, l_i^k\right]$. This implies that the demand for loans at bank i is given by $L^i = (l_i^j + l_i^k)L = \left[\frac{1}{n} + \frac{r_L^j + r_L^k - 2r_L^i}{2t}\right] L$. Hence, the bank's inverse loan demand function is $r_L^i = \frac{r_L^j + r_L^k}{2} + \frac{t}{n} - t\frac{L^i}{L}$.

Suppose that depositors also incur the proportional cost t with respect to distance from banks' attributes. Analogous reasoning then yields the supply of deposit funds to bank i, $D^i = \left[\frac{1}{n} + \frac{2r_D^i - r_D^j - r_D^k}{2t}\right] D$, an expression that can be rearranged in the form of an inverse supply function given by $r_D^i = t\frac{D^i}{D} - \frac{t}{n} + \frac{r_D^j + r_D^k}{2}$.

Given these market conditions and our maintained assumptions about bank balance-sheet and cost conditions (including $C_S \rightarrow 0$) profits of bank i are given by

$$\pi_i = \left[\frac{r_L^j + r_L^k}{2} + \frac{t}{n} - t\frac{L^i}{L}\right] L^i + r_S S^i - \left[t\frac{D^i}{D} - \frac{t}{n} + \frac{r_D^j + r_D^k}{2}\right] D^i - C(L^i, S^i, D^i).$$

All banks maximize profits with respect to the balance sheet constraint, $L^i + S^i = (1 - q)D^i$, which yields the profit-maximizing conditions

$$\left[-t\frac{L^i}{L}+\frac{t}{n}+\frac{r_L^j+r_L^k}{2}-\frac{t}{L}-C_L\right]=r_S,$$

and

$$-t\frac{D^i}{D}+\frac{t}{n}-\frac{r_D^j+r_L^k}{2}-\frac{t}{D}-C_D=(1-q)r_S$$

Because all banks are identical, in equilibrium loan and deposit rates much be the same, implying that ex post, $\frac{L^i}{L}=\frac{D^i}{D}=\frac{1}{n}$. Solving for the equilibrium loan and deposit rates yields

$$r_L=r_S+C_L+\frac{t}{L}$$

and

$$r_D=r_S-C_D-\frac{t}{D}$$

In this spatial-competition framework, the monopolistically competitive market loan rate equals the loan rate that would have arisen under perfect competition, r_S+C_L, plus a markup, $\frac{t}{L}$. The amount of the markup is increasing in the attribute cost per unit of total market loan demand. The monopsonistically competitive market deposit rate equals the perfectly competitive deposit rate, r_S-C_D, less a markdown, $\frac{t}{D}$. The magnitude of the markdown is decreasing in the attribute cost per unit of total market deposit supply. Thus, as the relative importance of attribute differentiation across banks shrinks in importance relative to the sizes of the loan and deposit markets, so do the magnitudes of the loan rate markup and deposit rate markdown, pushing these market rates closer to perfectly competitive levels. In contrast, if the cost associated with differences in attributes per unit of loans is relatively high—that is, as long as differentiation across banks' products and services is a significant factor in banking markets—then the market loan rate will reflect a larger markup, and the market deposit rate will reflect a larger markdown.

Summary: Models of the Banking Firm

- In general, perfectly competitive banks allocate assets and liabilities to satisfy joint marginal conditions across loans, securities, deposits, and nondeposit liabilities.
- Under certain conditions, including the key requirement that one balance-sheet item be tradeable at as either a net asset or a net liability at the same per-unit interest cost, portfolio separation holds, meaning that bank asset and liability decisions are mutually independent.

- Fundamental dynamics are important in banking if institutions face costs of adjusting loans or deposits, in which case contemporaneous lending and deposit funding choices depend on lagged choices as well as expected future interest rates.
- *Ceteris paribus*, perfectly competitive banking markets yield the highest feasible levels of bank lending and deposit funding, the lowest feasible loan rate and highest feasible deposit rate, and the greatest amount of consumer surplus. A bank loan-market monopolist restrains lending to maximize profits, resulting in less lending, a higher loan rate, a transfer of consumer surplus to the banking industry, and a deadweight loss of consumer surplus. A bank deposit-market monopsonist restrains deposit funding, which results in the issuance of fewer deposits, a lower deposit rate, a transfer of producer surplus to the banking industry, and a deadweight loss of producer surplus.
- Imperfect competition between the extremes of perfect competition and pure monopoly or pure monopsony includes either oligopoly or monopolistic/monopsonistic competition. In typical Cournot–Nash oligopoly models, banks to choose make their individual loan and deposit choices taking the choices of rivals offering identical loans and deposits as given. These oligopoly models yield equilibrium loan and deposit levels and rates between the extremes of perfect competition and monopoly/monopsony, with an increase in the number of banks pushing realized outcomes closer to those arising under perfect competition. Models of monopolistic or monopsonistic competition in banking allow for the possibility that banks' products are differentiated. Applying standard Chamberlin-type model of monopolistic/monopsonistic competition to banking markets also yields results approaching the limiting case of perfect competition as the numbers of banks in loan and deposit markets increase. In spatial models of monopolistic/monopsonistic competition in banking, costs incurred by consumers due to differences in banks' product or service attributes generate a market loan rate markup and deposit rate markdown in relation to perfectly competitive rates. Thus, reduced differentiation of banks pushes market loan and deposit rates toward perfectly competitive levels.

Chapter 4
The Industrial Economics of Banking

Policy-oriented economic analysis of the industrial structure of the banking industry traditionally has been guided by two intellectual paradigms: the structure-conduct-performance (SCP) hypothesis and the efficient-structure (ES) theory. The SCP hypothesis proposes that the level of concentration in a banking market influences banks' conduct, which in turn has a bearing on loan and deposit quantities, qualities, interest rates, and other market outcomes that determine consumer welfare. The SCP hypothesis implies that greater concentration gives banks more market power, which in turn leads to fewer loans and deposits and higher loan rates and lower deposit rates, all of which reduce consumer welfare. The ES theory suggests that cost conditions faced by banks play a crucial role in determining the optimal scale of individual banking organizations and the appropriate scope of banking activities. Thus, in contrast to the SCP hypothesis, the ES theory indicates that greater costs efficiencies resulting from expansions of scale and/or scope can lead both to expansions in loans and deposits, with associated lower loan rates and higher deposit rates. In practice, therefore, regulators contemplating applications for new banking licenses or, as discussed in the following chapter, proposed bank mergers and acquisitions have focused considerable attention on a perceived trade-off between resulting increases in market power versus cost-efficiency gains.

As we shall discuss shortly and as noted as well by Neuberger (1998), Shaffer (2004a), and Degryse and Ongena (2008), developments in the theory of industrial organization have added to the menu of factors that may influence bank market structure, conduct, and performance. Let's first begin our contemplation of the industrial organization of banking with an analysis of the fundamental implications of the traditional SCP theory.

The Structure-Conduct-Performance Paradigm in Banking

There have been several past expositions of the basic SCP paradigm as applied to the banking industry, including Gilbert (1984), VanHoose (1984), and Hannan (1991). We can readily illustrate the essence of the SCP paradigm as applied to banking

D. Van Hoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-642-02821-2_4, © Springer-Verlag Berlin Heidelberg 2010

markets by applying the basic oligopoly banking model discussed in the previous chapter.

The SCP Hypothesis with Identical Banks

Suppose that there are m loan-market rivals and n competitors in the markets for homogeneous loans and deposits. Banks are indistinguishable, and their marginal resource costs associated with loans (L), deposits (D), and nondeposit liabilities (N) have identical constant values. Chapter 3 showed that in this environment, a Cournot–Nash equilibrium yields the market loan rate $r_L = \Omega(r_N - C_N - C_L)$, where $\Omega = \frac{m}{m-\varepsilon^{-1}}$ and ε is the market loan demand elasticity. The market deposit rate under a Cournot–Nash equilibrium is $r_D = \theta[(1-q)(r_S - C_S) - C_D]$, where $\theta = \frac{n}{n+\eta^{-1}}$ and η is the market deposit supply elasticity.

The reasoning underlying the SCP hypothesis is depicted in Fig. 4.1, which revisits Fig. 3.7 in the previous chapter. In panel (a), if there are numerous loan-market competitors, so that $m \to \infty$ and $\Omega \to 1$, the market loan rate is at the perfectly competitive level r_L^{PC}, and borrowers' consumer surplus is the triangular region bounded beneath the market loan demand curve and the above the total marginal cost of lending, $r_N - C_N - C_L$. As the number of loan-market rivals decreases, however, so that $m \to 1$ and $\Omega \to \frac{1}{1-\varepsilon^{-1}}$, the monopoly markup, and the market loan rate rises toward the monopoly level, r_L^M. Industry economic profits expand toward the maximum, pure monopoly level equal to the area of the rectangle denoted B, which is a transfer from consumers to banks. Consumer surplus accruing to borrowers shrinks toward the area of the triangle denoted A above the monopoly loan rate r_L^M. Finally, because lending falls toward the monopoly level L_M as the number of loan-market rivals shrinks, a portion of the original consumer surplus, which when $m = 1$ equals the area of the triangle C in panel (a), becomes a maximized deadweight loss that is unattainable both to banks and their borrowers.

In panel (b)—not drawn to the same scale as panel (a)—if there is a large number of rivals in the deposit market, $n \to \infty$ and $\theta \to 1$, the market deposit rate is at the perfectly competitive level r_D^{PC}, and the amount of producer surplus accruing to depositors is the triangular region bounded above the market deposit supply curve and below the total marginal return on deposits, $(1-q)(r_S - C_S) - C_D$. Thus, as the number of competitors increases in the deposit market, $n \to 1$ and $\theta \to \frac{1}{1+\eta^{-1}}$, the monopsony markdown, and the market deposit rate declines toward the monopsony level, r_D^M. The total amount of monopsonistic exploitation of depositors enlarges toward the maximum, pure monopsony level equal to the area of rectangle F, which is a transfer from depositors to banks. Depositors' producer surplus shrinks toward the lower-left triangle denoted E. As the number of competitors in the deposit market shrinks, the amount of deposits at banks decreases toward the monopsony level D_M. A portion of the original producer surplus becomes the largest possible deadweight loss that is unattainable both to banks and their depositors, which when $n = 1$

reaches its maximum possible level given by the area of the upper right-hand triangle *G* in panel (b).

Clearly, the SCP hypothesis suggests that a reduction in the number of loan- and deposit-market competitors—that is, a rise in concentration in bank loan and deposit markets—generates imperfectly competitive conduct that results in higher market loan rates and lower market deposit rates. Loan and deposit quantities decline with greater concentration, which results in decreases in consumer and producer surpluses obtained by banks' customers and increases in deadweight losses. Consequently, increased concentration ultimately results in poorer market performance.

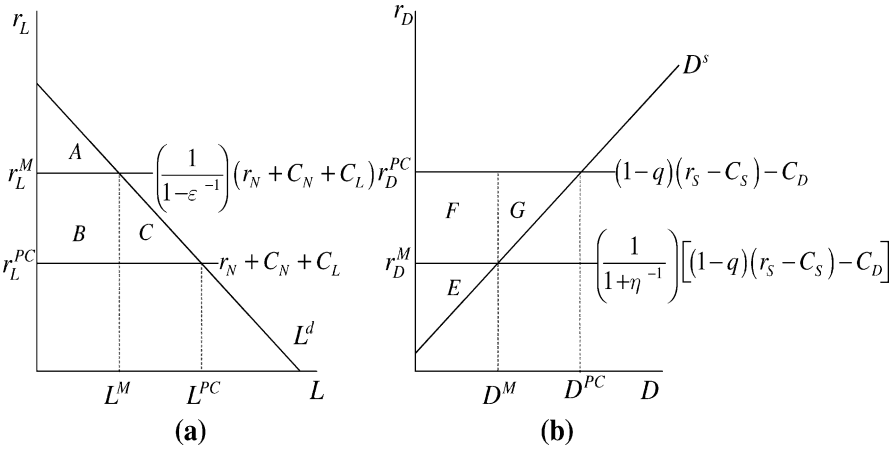

Fig. 4.1 The structure-conduct-performance hypothesis: a setting with identical banks

Structural Asymmetry, Dominant Banks, and the SCP Paradigm

Of course, a symmetric market environment such as the one considered above is unlikely to prevail in the real world. Instead, banks with differing costs compete side by side and operate at different scales. The identical-bank SCP hypothesis outlined above is obviously inconsistent with such a setting. Within the SCP paradigm, an alternative, dominant-bank model takes into the potential for more cost-efficient, larger banks to engage in market rivalry with less efficient, smaller banks.

A Dominant-Bank Model

Figure 4.2 provides an exposition of the dominant-bank framework as applied in a bank loan market. In this market, a large institution, bank *i*, and a number of smaller, "fringe," banks, indexed $j = 1,\ldots m$, raise funds by issuing nondeposit as well as deposit liabilities. These fringe banks are assumed to face costs of entering the market, so *m* is assumed to be a relatively small number. Both sets of institutions

face upward-sloping marginal resource costs. Bank i's total marginal cost of lending expressed in terms of funds raised via issuance of nondeposit liabilities is, as discussed in Chapter 3, given by $MC_L^i = \boldsymbol{r}_N + C_N^i + C_L^i$, which is the upward-sloping curve graphed in panel (a) of Fig. 4.2. The bank's average cost of lending is the curve labeled AC_L^i, which is assumed to correspond to the bank's long-run efficient scale of operations. If bank i were a monopoly bank, it would face the loan market demand curve L^d, its marginal revenue curve would be MR_M^i, and it would extend a profit-maximizing quantity of loans equal to L_M^i and charge the monopoly loan rate $r_L^{i,M}$. The bank's profits would equal the quantity of lending L_M^i multiplied by the difference between this loan rate and the value of AC_L^i evaluated at that amount of lending.

Each of the m fringe competitors in the loan market has access to less-efficient—compared with that available to bank i—technologies for raising and lending funds. For the sake of simplicity, suppose that each of these smaller banks thereby confronts a higher (at any given amount of loans) and identical total marginal cost of lending equal to $MC_L^j = \boldsymbol{r}_N + C_N^j + C_L^j$ shown in panel (b) of Fig. 4.2. As indicated, in the short run each of the identical fringe banks stands willing to lend as long as it earns a loan rate exceeding a reservation loan rate, $\bar{r}_L$—presumably a loan rate sufficiently low that individual banks cannot cover their variable costs. Furthermore, because each fringe bank j has access to intermediation technologies inferior to those possessed by bank i, it faces an average cost of lending, AC_L^j—assumed to correspond to each fringe bank's long-run minimum efficient operating scale—that lies above the average cost of lending faced by bank i for any given amount of lending. The amount of lending by each fringe bank is sufficiently small in relation to total market loans that no single fringe bank can affect the market loan rate. Thus, each of these smaller banks takes the loan rate as "given," and the total combined supply of loans by all of these smaller, fringe banks is, as depicted in panel (b), given by $L_F^s = \sum_{j=1}^{m} MC_L^j$.

In the face of competition from these numerous smaller institutions, above the fringe institutions reservation loan rate $\bar{r}_L$, bank i confronts not the entire market demand curve L^d in panel (a) but instead the *residual* demand curve $L_D^{i,d} = L^d - L_F^s$. That is, at each possible loan rate above $\bar{r}_L$, the total quantity of loans demanded by the public from the "dominant" institution, bank i, equals the market quantity demanded less the total amount supplied by all fringe banks at that loan rate. Associated with this residual demand curve is the marginal revenue curve MR_D^i. Bank i, therefore, maximizes its profits by extending a reduced amount of loans equal to L_D^i. The profit-maximizing loan rate is r_L^D, and bank i's maximum profits from lending drops to the quantity of lending L_D^i multiplied by the difference between r_L^D and the value of AC_L^i evaluated at L_D^i.

The fringe banks take the dominant bank's loan rate r_L^D as the given, market loan rate. Hence, the total quantity of loans they extend is equal to L_F in panel (b), which in turn must equal the total market quantity of loans demanded at the loan rate r_L^D less the amount of loans extended by the dominant bank i, L_D^i. In the situation

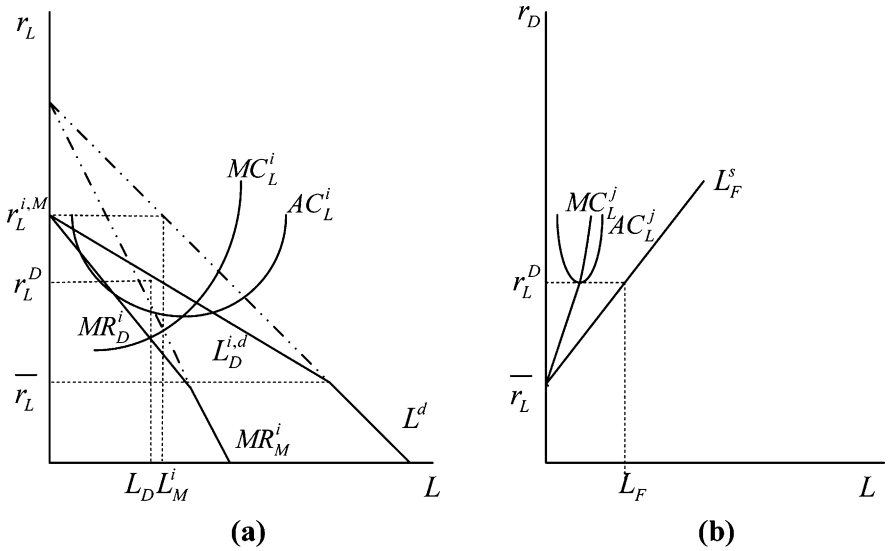

Fig. 4.2 The structure-conduct-performance hypothesis: a loan market with a dominant bank and a competitive fringe

depicted in panel (b), each fringe bank *j* earns zero economic profits, implying that in the specific situation depicted in the figure, there is no incentive for fringe banks either to enter or exit the market.

Thus, Fig. 4.2 illustrates a long-run equilibrium, prior to which there might have been fewer fringe firms in the market and hence a less elastic aggregate fringe supply curve, a higher profit-maximizing loan rate for the dominant bank, and positive profits for fringe firms that would have encouraged firm entry. Note that in the depicted long-run equilibrium, even though fringe banks earn no economic profits, the dominant bank is able earn positive (but non-maximized) profits due to its cost advantage.

Strategic Entry Deterrence

If fringe entry costs are relatively low, then a fundamental danger faced by bank *i* is that at some point in time following entry, one or more of its fringe rivals may, in the process of learning by doing, discover how to replicate the technology that provides bank *i* with its market edge. In a dynamic setting, recognition of this danger could give bank *i* an incentive to engage in strategic entry deterrence. Bank *i* might, for instance, engage in strategies aimed at raising the costs of its potential fringe rivals (see, for instance, Salop and Scheffman, 1983, 1987 and Scheffman and Higgins, 2003).

In the present context, however, bank *i* could utilize its existing technological edge to engage in predatory (or limit) pricing. That is, bank *i* could set its loan rate just below r_L^D in Fig. 4.2. As long as this loan rate exceeds its average cost

of lending at its profit-maximizing quantity of loans, bank *i* could thereby forestall entry by fringe banks and ensure itself a steady stream of profits equal to the shaded area. In a dynamic context, bank *i* would be more likely to opt for this strategy if its owners have a relatively low rate of time discount and if the bank has a sufficiently deep concern that a fringe rival could indeed learn by doing and replicate bank *i*'s intermediation technology. If so, the expected discounted stream of profits under this entry-deterrence strategy exceeds the discounted profit stream from pursuing a loan-pricing policy that permits entry. This strategy maximizes short-term profits at the expense of future economic profits.

Evaluating the Applicability of the SCP Paradigm to the Banking Industry

The SCP hypothesis suggests that in more concentrated banking markets, *ceteris paribus*, observed industry quantities of loans and deposits should be smaller. In addition, market loan rates should be higher, and market deposit rates should be lower. These quantity and interest-rate adjustments, of course, imply higher industry profits, reduced levels of consumer surplus received by borrowers in loan markets, and lower accruals to depositors of producer surplus in deposit markets.

Traditional SCP Evidence from Cross-Sectional Banking Data

Is there evidence that increased concentration generates these predicted effects within loan and deposit markets? Efforts to answer this question have traditionally focused on evaluating how empirical measures of bank performance have responded to changes in measures of bank market concentration. Not surprisingly, the performance measures most often utilized have been those for which data have been readily obtainable, including loan rates, deposit rates, and industry profits. Traditionally, the most common measure of bank market concentration have been one-, two-, or three-bank concentration ratios, or the combined assets, loans, or deposits of the top one, two, or three banks expressed as percentage shares of total market quantities. Another measure, less used in earlier studies but more commonly utilized in recent years, has been the Herfindahl–Hirschman index (HHI), which is the sum of squared percentage shares of assets, loans, or deposits of each of the banks in the market.

Gilbert (1984) provided an exhaustive review of numerous early studies evaluating the empirical relevance of the SCP hypothesis. These studies, which mostly examined U.S. banking data from the 1960s and 1970s, generally found that a 10-percentage-point increase in the concentration ratio was associated with an increase of average market loan rates of between 0.1 basis point and 11 basis points and a decrease in average deposit rates of between 0.1 basis point and 18 basis points. A 10-percentage-point rise in the concentration ratio was associated with an increase in banks' net income as a percentage of assets ranging between just over

1.5 percent and just under 9 percent and a rise in banks' net income as a percentage of equity capital ranging from less than 0.5 basis point to nearly 80 basis points. Although a few studies found responses of interest rates that were at odds with the SCP hypothesis or that were not statistically significant, the preponderance of evidence was generally supportive, within commonly accepted bounds of statistical significance, of key qualitative predictions of the SCP hypothesis: higher bank market concentration appeared to generate higher loan rates, lower deposit rates, and increased industry profits. Consistent with these predictions, Rhoades (1982) concluded from 1969–1978 U.S. banking data that monopoly power in banking reduced loans to individuals by about 16 percent and pushed up profits by approximately 13 percent.

There have been a number of evaluations of the SCP predictions for loan rates. Among more recent studies, Shaffer and Srinivasan (2002) have found strong evidence of higher loan rates resulting from increased concentration in data samples ranging from 2,500 to 3,900 U.S. banks. Mallett and Sen (2001) also conclude that there is a significant negative relationship between number of competing banks in Canada and small business loan rates. In addition, in a study of data from more than 200 Spanish banks in about 50 geographic markets, Martín-Oliver et al. (2008) find evidence that an increase in the number of banks in loan and deposit markets reduces market loan rates and boosts market deposit rates—although dispersion of loan and deposit rates also increases, which the authors attribute to search costs faced by consumers. Furthermore, in a study of more than 7,000 firms in a dozen nations between 1998 and 2005, Ongena and Popov (2009) find that increased integration of European interbank markets resulted in greater competition in credit markets, which resulted in lower market loan rates.

Among all specific studies of the SCP paradigm, one of the most influential has been Berger and Hannan's (1989) examination of concentration and performance in bank retail deposit markets. Berger and Hannan applied several different econometric techniques, a variety of empirical specifications, and a range of alternative measures of concentration to data from 470 U.S. banks between 1983 and 1985. They consistently found robust support for the basic predictions of the SCP hypothesis. More recently, Adams et al. (2002) utilize 1987–1996 U.S. banking data to estimate a framework based on a Cournot banking model and conclude that banks possess market power.

As Brewer and Jackson (2006) have noted, however, a key omitted variable in most empirical analyses of the SCP paradigm is risk. Suppose that, as discussed in the following chapter, greater concentration in banking markets has an impact on the risks incurred by banks. If so, loan and deposit rates would vary from perfectly competitive levels not as a result of imperfect competition alone but also as a consequence of endogenous adjustments of risk premia incorporated into these rates. When Brewer and Jackson account for risk in the derived demand for deposits, they find evidence that the effect of market concentration on deposit rates is about 50 percent smaller than corresponding estimates that fail to account for risk. Brewer and Jackson argue, therefore, that, traditional cross-sectional SCP studies likely overstate the influence of concentration on loan and deposit rates.

Evidence from Cross-Country Studies

Although much cross-sectional evidence in favor of the SCP hypothesis has focused on U.S. banking, a few recent studies have offered support for the central predictions of the SCP paradigm derived from international data. Shaffer (2001), for instance, examines banking data from 15 industrialized nations between 1979 and 1991 and finds evidence of market conduct consistent with predictions of Cournot oligopoly theory. In a study of 1981–1989 banking data from seven European nations, Neven and Röller (1999) find evidence supporting conduct consistent with a theory of cooperative, collusive behavior. In a study of data on banking markets in five EU countries in the mid-1990s, Goddard et al. (2004) conclude that there is a positive relationship between concentration and the traditional profitability performance measure.

Dynamic Interest Rate Responses: Competition and Pass-Through Effects

Since the early 1990s, a number of studies have broadened the scope of inquiry beyond cross-sectional analysis to consideration of time-series evidence relating loan and deposit rates to concentration. The focus of these studies has been the extent to which changes in interest rates determined in perfectly competitive markets for securities and other debt instruments pass through to loan and deposit rates in retail banking markets.

Most studies of interest-rate pass through suggest that the relationship between retail bank rates and interest rates on other financial instruments is an indicator of the degree of market competition among banking firms. Neumark and Sharpe (1992) and Sharpe (1997) argue that the intertemporal relationship among deposit rates and other market interest rates is hard to square with standard banking models without appeal to imperfect competition. Hannan and Liang (1993) have also considered the time-series relationship between market rates and bank retail rates in an effort to infer a relationship between bank market concentration and market deposit rates. Imperfect competition is a key feature in the analysis of Kahn et al. (1999), who focus on clustering effects in banking markets as possible factors influencing the determination of deposit rates, and Winker (1999), who suggests that asymmetric information and credit risks may also play some role in incomplete pass-through effects. In a recent examination of deposit-rate determination in U.S. banking markets, Rosen (2007) concludes that there is evidence of a complex dynamic interplay between the market competition and retail rates.

There are certainly theoretical grounds for the idea that the magnitude of pass-through effects should depend on the degree of competition in banking markets, as shown by Kopecky and VanHoose (2009). This study combines a linearized version of VanHoose's (1985) oligopoly model of bank loan and deposit markets with Elyasiani et al.'s (1995) model of fundamental dynamics discussed in the previous chapter. In the general case of imperfectly competitive loan and securities markets, the results are solutions for contemporaneous loan and deposit rates that depend on their own lagged values, lagged values of a competitive securities rate, and the

anticipated future values of the competitive securities rate. As a consequence, there is incomplete pass through from changes in the securities rate to bank loan and deposit rates. In perfectly competitive limits of this dynamic framework, however, the loan and security rates adjust one-for-one with the competitive securities rate, as predicted by standard static models of perfectly competitive banking markets.

In recent years, there has been an upsurge in studies of pass through from market rates to retail bank rates at individual institutions (Hofmann and Mizen, 2004); studies of specific countries such as Austria (Burgstaller, 2005), Canada (Scholnick, 1999), Chile (Berstein and Fuentes, 2004), Finland (Kauko, 2005), Germany (Winker, 1999), Malaysia and Singapore (Scholnick, 1996), Turkey (Aydin, 2007), the United Kingdom (Heffernan, 1997), and the United States (Hannan and Liang, 1993, Jackson, 1997, Kahn et al., 2005, Neumark and Sharpe, 1992, and Sheehan, 2006) and studies examining broader cross-country evidence, including Cottarelli and Kourelis (1994), de Bondt (2005), de Bondt et al. (2005), Égert et al. (2007), Espinoza-Vega and Rebucci (2003), Marotta (2007), Sander and Kleimeier (2004, 2006), Sørensen and Werner (2006), Tieman (2004), Toolsema et al. (2001), and van Leuvensteijn et al. (2008). These studies have reached mixed conclusions. Some studies, such as Neumark and Sharpe (1992) and Scholnick (1999) find evidence of asymmetric adjustment of bank retail rates to variations in market rates, but several others, such as Heffernan (1997), Hofmann and Mizen (2004), do not. Other studies, such as Jackson (1997) and Kauko (2005), suggest that nonlinearities appear to be present in the relationships between market rates and retail rates.

One finding common to virtually every study is evidence of sluggish and incomplete pass through from market rates to bank loan and deposit rates. In virtually every country examined, changes in market rates are not fully reflected in contemporaneous bank loan and deposit rates. Although the latest research finds evidence of slightly increased interest rate pass through to bank retail rates in some parts of the world in more recent years, cross-country variations remain as wide as first noted by Cottarelli and Kourelis (1994), who estimated contemporaneous impacts of market rates on bank rates varying from as low as 0.06 to not much above 0.80. Thus, time-series pass-through studies offer conclusions generally supportive of the SCP paradigm's predictions regarding the relationships between bank loan and deposit rates and other market rates.

The Conduct and Relative Performances of Large and Small Banks

The dominant-bank model utilized by proponents of the SCP paradigm offers a number of predictions. First, a situation with asymmetric competition depicted Fig. 4.2 cannot exist unless a dominant bank possesses a technological edge over its fringe rivals. Thus, in asymmetric banking-market environments we should observe large, dominant banks operating at lower costs than smaller rivals. As we discuss in the following section, there is indeed considerable evidence supporting the hypothesis that large banks tend to be more efficient than smaller competitors.

Second, if entry costs faced by fringe banks are substantial, then the existence of a cost advantage enables dominant banks to set their loan rate independently and

experience positive economic profits even when some fringe entry occurs. Costs of entry into banking markets may in fact be relatively significant, in which case economic profits accruing to dominant banks may be relatively substantial due to dominant banks' ability to restrain lending and mark up the loan rate over marginal cost. Nevertheless, if fringe entry can in fact take place at relatively low expense, then the dominant bank must respond by expanding its lending and reducing its loan-rate markup but may nonetheless reap positive profits on account of its technological advantage.

Consistent with the basic implications of the dominant-bank model, Craig and Hardee (2007) find evidence that in the market for U.S. small business loans, large banks are likely to extend less credit and that lending by smaller institutions fails to bring total market loans to competitive levels. Pilloff's (1999) findings are consistent with the prediction of the dominant-banking model in a setting with high fringe entry costs. Pilloff finds that the presence of very large and regionally prominent banks in a market tends to raise profitability both for those banks and for smaller competitors in that market. Richards et al. (2008) also finds evidence in banking 2005 data from upper-midwestern U.S. states consistent with considerable market power by dominant incumbents that he attributes particularly to spatial impediments to competition for borrowers. In addition, in a study of banking markets in 15 European Union nations between 1997 and 2004, De Jonghe and Vennet (2008) conclude that only larger banks are consistently able to earn non-competitive rents. In deposit markets, Hannan (2006), Hannan and Prager (2004, 1998), Park and Pennacchi (2008) find evidence of pricing behavior by large, multimarket banks and smaller, single-market banks that is consistent with application of the dominant-bank model to markets for deposit funds, although results reported by Rosen (2007) suggest that the impact of bank size and market structure on deposit rates may have dissipated more recently. Furthermore, Hannan and Prager (2009) conclude that the profitability of small banks in rural markets depends on whether these banks operate only in single markets and whether competing dominant and fringe banks are multi-market institutions. Finally, DeYoung (2003) finds evidence that exit rates of smaller banks are similar regardless of how long they have been present in a market, suggesting that the dominant-bank theory's treatment of fringe banks as similar institutions is reasonable.

Third, fringe banks that enter a dominant bank's loan market should charge the same loan rate as dominant banks. Coccorese (2009) provides evidence on bank pricing generally consistent with this prediction in a study of 86 single-market Italian banking institutions over the 1988–2005 interval. If fringe banks' loan rates lie above their average cost of lending, fringe banks should experience positive economic profits. Nevertheless, in light of both the smaller scale of lending by a fringe bank and the higher average cost it faces, fringe bank's profitability should be lower than the profitability of dominant banks. Consistent with these predictions, Amel and Liang (1997), who study data from more than 2,000 U.S. banking markets from the late 1970s through the late 1980s, find evidence that bank entry responds to positive profits. They conclude that profits of incumbent large and particularly small banks decline in response to such entry.

Fourth, in principle a dominant bank can forestall entry by fringe competitors by setting its loan rate lower than the minimum long-run average cost of potential fringe entrants. This reduces—possibly substantially—the dominant bank's short-run profits but in theory could enable the bank to maintain the technological edge that it possesses over potential entrants. In this way, the bank might be able to maintain a steady stream of positive economic profits over a longer horizon, which is consistent with evidence provided by Shaffer's (2002) study of a monopoly bank in a small Texas community between 1984 and 1999.

Not all empirical evidence is completely supportive of the dominant-bank model's predictions. Strahan (2008) for instance, concludes that large banks may be more likely to lend than small banks. Berger et al. (2007) find that when they utilize a measure of market size structure takes into account assets both within and outside a given market, no clear advantage accrues to large banks over small banks in the market for small business lending. In addition, Gobbi and Lotti (2004) find evidence from data for a 1990–2002 entry liberalization period in Italian banking markets that entry by de novo banks is more likely in response to profit opportunities than the opening new branches by established banks, which they hypothesize results from possession of information by local financial entrepreneurs not available to large incumbent banks. Notably, however, Felici and Pagnini (2008) find in an examination of Italian banking data over nearly same interval that distance-related entry costs present greater barriers to entry for small banks than large banks. Separately, Bikker and Haaf (2002) apply the approach to measuring bank-level competition developed by Panzar and Rosse (1987), which involves utilizing an index reflecting the extent to which an increase in input prices are reflected in revenues. Bikker and Haaf utilize 1990s data from more than 5,000 banks in 23 countries and conclude that markets containing both large and small banks are monopolistically competitive. They find evidence that, if anything, smaller banks typically operate in less competitive markets—presumably more often in rural areas—than do larger banks. Brissimis and Delis (2009) likewise apply the Panzar–Rosse methodology. They examine data from 465 banks in 20 emerging economies and conclude that market power held by individual banks varies considerably both across countries and, potentially consistent with the dominant-bank model, within countries.

Market Structure and Bank–Customer Relationships

Recently, Boot and Marinč (2008) have suggested that relationship banking is a "prime source of banks' comparative advantage." Indeed, banking scholars have long emphasized the importance of customer relationships. In a survey of potential barriers to entry in banking that built on work of Bain (1956) in the SCP tradition, Alhadeff (1974) included advantages possessed by incumbents such as a leverage—that is, capitalization in relation to assets—advantage and an absolute cost advantage. In addition, he includes a product differentiation barrier and

argues that a key determinant of the height of this entry barrier is the bank–customer relationship. This relationship, which also received considerable attention in early work by Hodgman (1963), arises in large part from asymmetric information about risk and associated costs. More recently, in a context in which the relevant "customer" is a firm, Petersen and Rajan (1995) refer to the bank–customer relationship as "close and continued interaction" that "may provide a lender with sufficient information about, and a voice in, the firm's affairs."

Basic Market-Structure Implications of Bank–Customer Relationships

Wood (1975) provided the first analysis of fundamental behavioral implications of bank–customer relationships. Consider the following SCP-style elaboration of Wood's analytical framework. An imperfectly competitive bank i that has an established set of relationships with borrowers and depositors faces the loan demand function for period t given by $L_t^{i,d}(r_{L,t}^i;L_{t-1}^i,L_{t-2}^i,...L_{t-m}^i;X_t^i)$, and confronts the deposit supply function $D_t^{i,s}(r_{D,t}^i;D_{t-1}^i,D_{t-2}^i,...D_{t-m}^i;Y_t^i)$, where $\frac{\partial L_t^{i,d}}{\partial r_{L,t}} < 0$; $0 < \frac{\partial L_t^{i,d}}{\partial L_{t-j}^i} < 1$, $j = 1,\ldots, m$; $\frac{\partial D_t^{i,s}}{\partial r_{D,t}} > 0$; $0 < \frac{\partial D_t^{i,d}}{\partial D_{t-k}^i} < 1$, $k = 1,\ldots, n$; and X_t^i and Y_t^i are other factors influencing the demand for loans and the supply of deposits, respectively, that this bank faces, including perhaps decision variables of industry rivals that are taken as given in this analysis. Thus, the quantity of loans demanded by customers of an imperfectly competitive bank during the current depends in part on quantities of loans received from the bank in previous periods, and the contemporaneous amount of deposits supplied to the bank depends directly on lagged customer deposit holdings.

To simplify analysis of this version of Wood's framework, consider a two-period setting, so that $t = 2$ and $t–1 = 1$, and suppose that $m = n = 1$. Thus, in period 1, the bank maximizes the period 1 expectation of the discounted present value of profits over the two periods, given by

$$E_1\,(V) = r_{L,1}^i L_1^i + r_{S,1} S_1^i - r_{D,1}^i D_1^i + R[r_{L,2}^i L_2^i + r_{S,2} S_2^i - r_{D,2}^i D_2^i],$$

where R is the bank's subjective discount factor, with $0 < R < 1$, and, to further simplify, resource costs are assumed to be negligible (implying portfolio separation, as discussed in Chapter 3). Finally, suppose that the bank chooses its loan rate, securities holdings, and deposit rate to maximize V subject to the balance sheet constraint in each period specifying that loans plus securities must equal deposits net of reserve requirements, where q is the required reserve ratio. This implies that in the first period, the bank's choices must satisfy the following conditions:

$$\left[1-(\varepsilon^i)^{-1}\right] r^i_{L,1} = r_{S,1} - R\frac{\partial L^{i,d}_2}{\partial L^i_1} E_1\left(r^i_{L,2} - r_{S,2}\right), \text{ and } r^i_{D,1}\left[1+(\eta^i)^{-1}\right]$$
$$= (1-q)r_{S,1} + R\frac{\partial D^{i,s}_2}{\partial D^i_1} E_1\left[(1-q)r_{S,2} - r_{D,2}\right],$$

where, as in Chapter 3, ε^i is the loan demand elasticity and η^i is the deposit supply elasticity.

Note that if $\frac{\partial L^{i,d}_2}{\partial L^i_1} = 0$ and $\frac{\partial D^{i,d}_2}{\partial D^i_1} = 0$, so that there are no holdover effects from bank–customer relationships going into the second period, the above profit-maximizing conditions for first-period choices reduce to conditions analogous to those discussed in Chapter 3. If bank–customer relationships are present, however, $0 < \frac{\partial L^{i,d}_2}{\partial L^i_1} < 1$ and $0 < \frac{\partial D^{i,d}_2}{\partial D^i_1} < 1$. As a consequence, the terms $R\frac{\partial L^{i,d}_2}{\partial L^i_1} E\left(r^i_{L,2} - r_{S,2}\right)$ and $R\frac{\partial D^{i,s}_2}{\partial D^i_1} E_1\left[(1-q)r_{S,2} - r_{D,2}\right]$ appear as what Wood refers to as "imputed values of patronage" arising from the interperiod loan demand and deposit supply interactions arising from the existence of bank–customer relationships.

The inclusion of these imputed values implies that the bank will set a lower loan rate and higher deposit rate in the first period than it would have established in the absence of relationships with its borrowers and depositors. Doing so reduces the bank's profits slightly in the first period but enables it to attract more loans and deposits in the first period, which through its customer relationships will, *ceteris paribus*, generate higher loan demand and deposit supply in the second period. Thus, in the second period, the bank can set a *higher* loan rate and a *lower* deposit rate than it otherwise would have established in the second period, thereby boosting its second-period profits. Extending the above model to a longer horizon would yield the prediction that during the early periods of relationships with borrowers and depositors, loan rates would be lower, and deposit rates would be higher. In later periods, however, loan rates would increase, an deposit rates would drop.

Naturally, the lower the bank's subjective rate of time discount and hence the higher the discount factor, the greater will be the imputed benefits arising from bank–customer relationship. Hence, the bank will be more willing to trade off fewer profits in the present for greater profits in the future. Furthermore, the ability of the bank to mark up its loan rate and to mark down its deposit in the future declines with higher values of the loan demand and deposit supply elasticities, so an increase in competition reduces the present value of the bank's relationships with its borrowers and depositors.

Consistent with Alhadeff's argument, therefore, Wood's framework suggests that in an imperfectly competitive banking industry, contemporaneous loan-rate markups will be smaller and deposit-rate markdowns will be greater at incumbent banks with pre-established customer relationships than would be consistent with profit maximization on the part of a new entrant to the banking industry possessing no such customer relationships. Current profits of incumbents will be lower, reducing the

incentive for entry to occur and resulting in a more concentrated banking industry, *ceteris paribus*. Of course, incumbents have an incentive to utilize their market power to boost loan-rate markups and increase deposit-rate markdowns over time, which ultimately could lead to more entry. Nevertheless, Wood's model indicates that bank–customer relationships tend to forestall near-term entry.

Evidence on Bank–Customer Relationships

Wood's analysis simply takes as given the existence of bank–customer relationships that generate intertemporal increases in credit demand and deposit supply. It does not explain why customer relationships exist.

Determinants and Impacts of Bank–Customer Relationships

Virtually all work on bank–customer relationships, surveyed in depth by Boot (2000), has focused on relationships between banks and business borrowers. Berger and Udell (2002) identify three fundamental aspects of relationship lending from a bank's point of view: (1) a dependence on "soft" information not readily observed, verified, or transmitted about a firm, its owner, and its regional market; (2) coordination of this information by a bank loan officer; and (3) a range of contracting problems involving the borrower, loan officer, senior management, owners, and regulators that the bank must resolve. As a consequence, Berger and Udell conclude, smaller banks with fewer management layers are likely to have a comparative advantage in relationship lending. DeYoung (2008) suggests that this comparative advantage has fueled general bifurcation of the banking industry into a group of mainly smaller institutions emphasizing relationship banking involving lower-volume, high-value-added, personalized services and a separate group of primarily larger institutions specializing in the high-volume provision of relatively standardized and low-cost services and a separate

By emphasizing how relationships with customers boost the demand for loans and supply of deposits confronted by a bank in future periods, Wood's analysis abstracts from specific factors that might help motivate bank–customer relationships. Sharpe (1990) has offered a theory of bank–customer relationships in which banks develop reputations through implicit contracts. Maintaining these contracts enables banks to acquire private information about borrowers and thereby take advantage of captive customers in order to earn rents. In contrast, Blackwell and Winters (1997) have suggested that borrower relationships are valuable to banks because such relationships permit banks to reduce monitoring and hence reduce their costs. The result is lower loan rates to customers with which banks have relationships. Blackwell and Winters examined a sample of 174 business lines of credit with six banks and find both that banks reviewed less frequently firms with which they had longer relationships and also charged these firms lower loan rates.

Hannan (2008) explores how the expected duration of relationships with depositors influences the deposit rates paid by banks. In a study of more than 13,000

U.S. banks over the 1989–2006 period, Hannan finds that banks tend to offer higher deposit rates within areas and over times in which there was greater market in-migration, presumably reflecting a greater incentive to offer more appealing deposit rates in an effort to attract new depositors into relationships. In contrast, in areas and times in which there was greater out-migration, Hannan finds, consistent with the hypothesis that banks in such markets anticipate less durable relationships, evidence of lower deposit rates. Carbo-Valverde et al. (2009) reach analogous conclusions in a study of 65 Spanish banking institutions between 1986 and 2003.

Broecker (1990) and Dell'Ariccia et al. (1999) motivate relationship lending as a barrier to entry, but these authors go beyond Alhadeff's assertion that such a barrier simply exists to describe how it can arise endogenously in banking markets. In Broecker's framework, banks offer credit to borrowers that pass binary creditworthiness tests. As more banks enter the market, the average creditworthiness of firms that pass at least one test is decreasing, which results in adverse effects on equilibrium interest rates. The result is higher costs for banks, which drives up costs faced by potential entrants. Dell'Ariccia et al. (1999) provide a more explicitly relationship-based analysis in which new entrants to bank credit markets contemplating loan applicants must overcome adverse selection problems that have already overcome by incumbent banks that maintain relationships with successful borrowers, which gives the latter an informational advantage that constitutes an entry barrier.

Presumably relationships with banks can also be valuable to borrowers, and Petersen and Rajan (1994) find evidence of relationship value for small U.S. businesses in the late 1980s. Carletti (2004) develops a theoretical framework in which a firm seeking bank credit to fund a project balances monitoring costs imposed by multiple banks with benefits in the form of higher anticipated return from the project as a result of monitoring by multiple banks. She also examines the optimal monitoring choices of banks and identifies credit-market equilibrium outcomes in which lending relationships exist. Berlin and Mester (1999) additionally argue that an important feature of relationship banking is access to inelastically supplied core transaction and savings deposits that enable a bank to provide borrowers with insurance against exogenous shocks. Their study of a panel of U.S. banks between 1977 and 1989 finds evidence supporting this hypothesis.

Akhavein, et al. (2004) provide support for Alhadeff's suggestion that relationship lending acts as a barrier to entry. Akhavein et al. examine data on loans to farms by rural U.S. banks between 1987 and 1994, and they find that the length of tenure of a farm operation leads to more lending to the farm on the part of incumbents and less lending by de novo banks. Hence, they conclude that new entrants to banking markets during this interval do appear to have been placed at a competitive disadvantage by relationship lending.

Wood's framework predicts that over earlier periods of a lending relationship, loan rates paid by borrowers should be lower, consistent with Berger and Udell's (1995) finding, in a study of more than 3,000 business borrowers in the late 1980s, that borrowers involved in lending relationships do pay lower loan rates. Wood's model also indicates that a key value of relationship lending is the capability for lenders eventually to charge higher loan rates to borrowers with whom they have

maintained a continuing relationship. Degryse and Ongena (2005) study nearly 18,000 loans extended by a Belgian bank in the mid-1990s and find that the loan rates that the bank charged borrowers did indeed increase with the duration of the bank's lending relationships. They also find evidence that physical distance between lenders and borrowers helps to explain loan rates, which they suggest provides evidence of spatial price discrimination on the part of banks.

In a study more than 25,000 small-business loans extended by a U.S. bank located in New England, however, Agarwal and Hauswald (2007) find that once the bank's proprietary and private information is taken into account, distance effects on loan rates disappear, suggesting that the main effect of distance is its impact on banks' capabilities to acquire and take advantage of "soft" information in relationships with borrowers. Ratti et al. (2008) find evidence from data for non-financial firms in 14 European nations for the 1992–2005 interval that financial constraints are less binding on firms when banking markets are more concentrated, which they suggest indicates that banks with market power face lower costs of acquiring and acting on information regarding the creditworthiness of potential borrowers.

Ongena and Smith (2001) examine data from Norwegian business borrowers between 1979 and 1995 to identify factors influencing the duration of lending relationships. They find that firms that are smaller and more profitable tend to maintain shorter relationships. So do firms that borrow from multiple lenders.

Competition and Relationship Lending

How does increased competition in loan markets affect relationship lending? On the one hand, based on a model that focuses on the capability of firms to internalize benefits of lending relationships, Petersen and Rajan (1995) argue that the benefits of lending relationships dissipate with greater competition in the loan market. This is a prediction that is also consistent with Wood's framework. Von Rheinbaben and Ruckes (2004) reach a similar conclusion in a model that focuses on the potential transmission of proprietary information.

On the other hand, Boot and Thakor (2000) contend that in fact, more competition in *capital* markets should reduce relationship lending by banks. Their argument is that relationship lending increases the borrower's probability of success and hence a lender's return, so a bank responds to greater competition in the *loan* market by *expanding* its relationship lending. As banks do so, the marginal benefit from maintaining a borrower relationship declines. Yafeh and Yosha (2001) reach an analogous conclusion. Separately, a model of bank commitment to a borrower developed by Dinç (2000) predicts that increased competition in the loan market reinforces a bank's incentive to maintain a relationship with a borrower only if there initially are a small number of banks. Likewise, the theoretical analysis of Anand and Galetovic (2006) suggests that effects of greater competition on relationships in banking markets are potentially ambiguous.

The evidence regarding the direction of the effect of greater loan-market competition on relationship lending is mixed. Utilizing data on more than 3,000 U.S.

business borrowers in the late 1980s, Petersen and Rajan examine the relationship between loan market concentration and the portion of firm debt financed institutionally. They find that decreased concentration among lenders is associated with younger firms—hence firms that have less scope for establishing lending relationships—financing a larger share of their indebtedness with bank loans obtained at lower rates of interest. Consistent with Petersen and Rajan, in an analysis of 2003–2005 data on U.S. small business lending, Laderman (2008, 2007) concludes that in markets with larger proportions of young firms, greater concentration boosts the volume of small business loans. Zarutskie (2006) also finds evidence supporting Petersen and Rajan in her study of the impact of the U.S. Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994, which she finds that newly formed firms faced toughened borrowing constraints when U.S. bank competition increased.

In an analysis of loans granted by a Belgian bank to more than 13,000 firms in the mid-1990s, however, Degryse and Ongena (2007) reach the conflicting conclusion that *reduced* branch-banking concentration was associated with *more* relationship lending. In a study of borrowing by more than 4,000 Italian firms, Presbitero and Zazzaro (2009) suggest that these results may be driven partly by size and distance effects rather than by the level of bank market concentration, per se.

Furthermore, an examination of the effects of small business lending on U.S. bank performance between 1996 and 2002 by Ergungor (2005) indicates competitive pressures on community banks have reduced the profitability of relationship lending, thereby providing these banks with a reduced incentive to continue relationships with borrowers. Fields et al. (2006) likewise conclude, based on analysis of U.S. bank loan announcements from 1980 to 2003, that as competition among lenders increased over time the abnormal returns associated with loan announcements first noted by James (1987) dissipated, suggesting that advantages of relationship lending diminished. Furthermore, consideration of self-reported relationships by 122 German banks leads Elsas (2005) to conclude that in highly concentrated markets, less competition fosters relationship lending. Finally, de la Torre et al. (2008) conclude that the recent intensification of competition in lending to smaller businesses in many developing nations has occurred in the absence of relationships. Thus, so far the weight of the evidence appears to support the conclusion that relationship lending has declined in the face of increased loan-market competition.

The Efficient Structure Theory and Banking Costs

The SCP paradigm's dominant-bank model relies on the assumption that large banks have cost advantages over smaller rivals. A potential source of such an advantage is lower average per-unit operating costs accompanying an expansion in a bank's assets, or its scale. Alternatively—or perhaps additionally—lower per-unit costs of bank operations could possibly result from a broadening of the bank's product mix, or scope.

The Efficient Structure Challenge to the SCP Paradigm

The presence of significant economies of scale and scope could have implications for the relationship between bank market structure, conduct, and performance. Recognition of this fact forms the basis for the efficient structure theory, which suggests that scale and scope economies could account for the existence of relatively large banking organizations.

In contrast to the SCP hypothesis, the efficient structure theory proposes that a consequence of cost advantages due to scale or scope is *lower*, rather than higher, loan rates and *higher*, instead of lower deposit rates. Indeed, the efficient structure theory inverts key predictions of the SCP dominant-bank theory. The interest-rate settings of larger banks experiencing lower per-unit costs constrain the rates charged by smaller fringe rivals and thereby yield lower average loan-market rates and higher average deposit-market rates. As a consequence, the efficient structure theory predicts no clear relationship between market concentration and loan and deposit rates.

Building on Demsetz (1973), the efficient structure theory suggests that, although profit-rate differentials across banks should be driven down by competition in banking markets, at any given point in time larger institutions could operate with higher profits than smaller banks. According to the theory, higher profits observed at larger institutions would result from the efficiency advantages they possess rather than from predatory conduct aimed at precluding entry.

Furthermore, the efficient structure theory suggests that the intensity of market competition and market concentration are not necessarily negatively related. Indeed, in an application of the Panzar–Rosse methodology for measuring competition to data from more than 4,000 banks in 50 nations, Claessens and Laeven (2004) find no evidence that the Panzar–Rosse index of competition is related to banking industry concentration. They conclude that governmental entry and activity restrictions on banks play a much more significant role in affecting the observed level of competition.

Banking Efficiency and Costs

Evaluating the technical efficiency of the banking industry has been the subject of a considerable amount of work. A recent survey by Berger (2003) indicates that concentrated attention to new developments in information and financial technologies have contributed to improved technical efficiency in banking. Furthermore, Alam (2001) provides evidence that the major contributor to cost efficiency improvements in the U.S. banking industry during the 1980s derived primarily from technological change rather than changes in output scale or convergence to an efficient production frontier.

Consistent with Alam's findings, Berger and Humphrey (1991) suggest that the main source of cost inefficiencies derived from failure to utilize the least-cost production technology or the least-cost mix of inputs—often called X-inefficiencies—in banking are technical inefficiencies rather than inefficiencies in scale or product

mix. Berger and Mester (1997) conclude that a number of additional factors account for inefficiencies in banking, including organizational form, market characteristics, and regulation. Some have suggested that banking institutions might be more likely to exhibit X-inefficiencies because a number are either mutual institutions (about 8 percent in the United States) or not publicly traded (nearly 90 percent in the United States). Altunbas et al. (2001), however, find no evidence of systematic differences in cost efficiency across mutual, private, and publicly traded banking institutions, although Berger and Hannan (1998) argue that cost inefficiencies resulting from absence of market discipline in banking may create a social loss exceeding by several times the standard deadweight loss owing to monopoly power.

Wheelock and Wilson (1999) conclude that cost inefficiencies at U.S. bank between 1984 and 1993 arose mainly from failures of banks to adopt technological improvements implemented by large banks operating along the efficient frontier. There is evidence that sources of inefficiency vary across bank product lines, which could help explain why Wheelock and Wilson find evidence of greater technical efficiency at large banks. Devaney and Weber (2002), for instance, find that technical inefficiencies in small-business lending are much smaller than allocative efficiencies resulting from failure to utilize the least-cost mix of inputs. Sources of relative technical efficiency levels also appear to differ between large and small banks, an observation that DeYoung et al. (2004) argue is consistent with small banks specializing in using "soft," more qualitative information in making nonstandardized loans and large banks specializing in utilizing "hard," easily quantifiable information to extend relatively standardized loans. Carter and McNulty (2005) find evidence supporting DeYoung et al.'s hypothesis in U.S. banking data over the 1993–2001 period. Carter and McNulty conclude that smaller banks exhibit better performance in market for small business lending and larger banks performing better in market for credit card lending.

Another factor that helps to explain differences in efficiency levels for large versus small banks may be that larger banks engage in more off-balance-sheet activities. Clark and Siems (2002) argue that accounting for off-balance-sheet activities is important in estimating banking X-inefficiencies. Once they take such activities into account, Clark and Siems find that average profit efficiency across all banks is 25–35 percent lower than at the best-practice banks, or slightly above the 20–25 percent cost savings that Mester (2008) concludes banks could attain by improving their technologies or mix of inputs.

It is important to note, however, that researchers continue to have disagreements in interpreting studies of X-inefficiencies. As noted by DeYoung (1998), it may be that higher-quality management requires expenditures not made at less-well-managed banks. According to this view, "best-practice banks" may not necessarily be the banks with the lowest recorded costs.

Irrespective of how efficient banks may be technically or with respect to their input choices, is there evidence regarding scale or scope economies that supports the efficient structure theory? The answer to this question is "perhaps." Early studies of the U.S. banking industry conducted by Benston et al. (1982), Clark (1984), Gilligan and Smirlock (1984), and Gilligan et al. (1984) found little evidence of significant

scale economies in banking. Hunter and Timme (1995) reached the same conclusion after taking into account the quasi-fixed nature of physical capital and retail deposits. Berger et al. (1987) concluded that if anything banks experienced slight diseconomies of scale, and Gilligan and Smirlock likewise found evidence of diseconomies of scale for large banks. In a study of banking data for 15 nations over the 1988–1992 period, Allen and Rai (1996) also conclude that large banks experience scale diseconomies and find evidence of slight scale economies at small institutions. Allen and Rai also conclude, however, that output inefficiencies are dwarfed by X-inefficiencies. The bulk of these studies relied on estimates derived from translog cost functions, which Shaffer (1998) argues bias empirical results toward favoring a finding of economies of scale, hence suggesting even more strongly a general paucity of strong evidence of significant scale economies in most studies.

Hughes and Mester (1998) and Hughes et al. (2001) have argued, however, that incorporating capital structure and risk into bank scale economy measurement might lead to a greater potential for scale economies in banking. Hughes and Mester offer evidence from U.S. banking data from 1989 and 1990 that this is indeed the case, and Hughes et al. (2001) also find evidence of significant scale economies in a detailed analysis of 1994 U.S. banking data. Bossone and Lee (2004) apply the approach laid out by Hughes and Mester and Hughes et al. to data from 875 banks located in 75 different nations and also conclude that banking is characterized by significant scale economies once financial capital structure and risk are taken into account.

Evidence on scope economies is even more mixed. Berger et al. (1987) also argued that banks face slight diseconomies of scope, but in contrast Gilligan and Smirlock (1984), and Gilligan et al. (1984) concluded that the evidence at that time suggested at least slight jointness in banking production. Nevertheless, Berger et al. (1996) conclude that there was no evidence of scope economies revealed by U.S. banking data between the late 1970s and 1990. In a study of the determinants of U.S. banks' return on assets between 2000 and 2005, Asaftei (2008) finds that a key benefit of variations in product mix is to provide flexibility in sources of bank earnings. Asaftei finds, in fact, that the revenues forthcoming from a broadening of product mix helped to offset increases in costs during this period.

Efficient Structure Theory and Bank Performance

The mixed evidence on relative cost efficiencies across the size distribution of banks and on scale and scope economies yields a uncertain verdict on the relevance of the efficient structure theory for the banking industry. Much of the work favoring the efficient structure theory has focused on measures of bank performance.

In a study of U.S. banking data for the 1980s, Berger (1995) explores the relationship between profits and market structure within an empirical framework that aims to takes into account X-inefficiencies. On the one hand, Berger finds support for the efficient structure theory's prediction that higher profits accrue to the more

cost-efficient banks. On the other hand, he also finds some evidence in support of the SCP dominant-bank model's prediction that large banks are able to boost their profits through exercise of market power as well as cost-efficiency advantages. The latter result is consistent with the conclusions of Berger and Mester (2003), who have found that between 1991 and 1997, revenue production efficiency increased at U.S. banks, but cost efficiency did not.

Jayaratne and Strahan (1998) examine U.S. banking data stretching from the mid-1970s to the early 1990s, a period in which bank branching was deregulated by many states and bank costs declined as a result. Consistent with the efficient structure theory, they find that most cost-efficient banks grew larger at the expense of rivals that tarried in finding cost efficiencies during the period. Furthermore, they conclude that decreases in banks' costs contributed to lower market loan rates.

Several other studies offer performance evidence inconsistent with the SCP hypothesis. Smirlock (1985) for instance, finds that bank profits are related to market share instead of market concentration, consistent with the efficient structure theory's prediction that more efficient banks grow and prosper at the expense of less efficient banks. As Gilbert (1984) notes, Glassman and Rhoades (1980) also report empirical results for the same 1970s banking period that are similar to Smirlock's findings. Smirlock and Brown (1986) suggest that the greater efficiency of larger banks allows them to act as dominant market participants along lines of the SCP dominant-bank model but that observed higher profits earned by the larger banks result from greater efficiency rather than market power. Consistent with this perspective, Calem and Carlino (1991) examine interest rates on money market deposit accounts and 3- and 6-month certificates of deposits at 466 banks in 148 metropolitan statistical areas and conclude that banks behave strategically irrespective of market concentration. In an analysis of 1992–1999 European banking data, De Guevara et al. (2005) find little evidence that concentration affects bank deposit rates in European markets. Allen et al. (1991) study of fee dispersion across banks in the early 1980s, conclude that there is no evidence to support the SCP paradigm's prediction that market should explain variations in bank fees. Instead, they find that asymmetric information is the main explanatory variable.

Allen et al. (1991) is one of a set of studies applying game-theoretic-based models to banking. In a conjectural-variations analysis that seeks to account for oligopolistic interdependence among more than 170 Norwegian banks in 1988, Berg and Kim (1994) find that they can reject the Cournot behavior typically assumed by the SCP paradigm. They also find that accounting for oligopolistic interdependence leads to greater evidence favoring scale economies in banking. In a separate contribution, Berg and Kim (1998) apply a conjectural-variations approach to econometric analysis of a panel of data from Norwegian banks during the early 1990s to test whether these banks made choices taking into account rivals' responses. Berg and Kim find evidence of strategic interdependence in retail loan markets but relatively competitive behavior in the corporate loan market; yet, they reject Cournot behavior in both. Shaffer (1993, 1989) also utilizes a conjectural-variations methodology in analyzing Canadian banking data for the 1965–1989 period and U.S. banking

data between 1941 and 1983 and concludes that both nations' banking industries exhibited perfectly competitive behavior.

Based on an analysis of European banking data from the latter 1990s Corvoisier and Gropp (2002) also reject a Cournot model of bank loan and deposit market behavior. They find evidence in European banking data of efficiency and contestability—competitive behavior of incumbents in light of threats of potential entry—irrespective of bank market concentration, a conclusion that is consistent with Weill (2008)'s finding of convergence of banking efficiency in European nations between 1994 and 2005.

In some cases, authors conclude that behavior in certain banking markets entails Bertrand–Nash behavior involving price competition among oligopolistic rivals offering heterogeneous products. Barros (1999) applies a Salop-style model (see Chapter 3) to develop testable implications with respect to Bertrand–Nash versus collusive behavior and considers data from 15 Portuguese banks in the early 1990s. He concludes that there is no evidence of market-wide collusion and that when spatial transport costs are taken into account Bertrand–Nash behavior receives greater support from the data. In addition, Molnár (2008) develops a framework that tests market power in the Finnish banking industry by deriving price-cost margins predicted by alternative strategic oligopoly models. He seeks to determine the theory that best fits the actual market environment by matching actual data on cost-price margins to the predicted margins. Molnár concludes that Finnish banks also exhibit Bertrand behavior. In eight-firm-dominated Italian banking market, Coccorese (2005) concludes that conduct is even more competitive than outcomes predicted by the Bertrand–Nash model of oligopoly rivalry.

Endogenous Sunk Fixed Costs and Banking Industry Structure

Most analyses of banking cost efficiency assume that a bank's fundamental objective is to attain a cost-minimizing scale and scope of operations. Given attainment of this minimum-cost objective, the bank determines its balance sheet with an aim to maximize profits. According to this perspective, the only sunk costs of banks are exogenous costs of entry and setup. A predictable consequence of exogenous sunk costs is as that the size of a competitive banking market expands, perhaps through growth in income or population, incumbent banks should reach their efficient scales. As profits increase, additional banks should enter and should as well pursue output expansions toward efficient scales. Thus, as market size expands, there should be an increase in the number of banks and an accompanying decline in each bank's market share. Consequently, an increase in market size eventually should lead to a reduction in concentration.

Sutton (1991) proposes, however, that in some market environments, sunk costs may be *endogenous*. An important consequence, Sutton suggests, is that industry structure may remain static even in a growing market. Firms in such industries, Sutton suggests, select a stream of fixed outlays on items such as research and development, advertising, or other characteristics that enhance the demands for

their products. As a result, an expansion in market size encourages firms to proportionately increase their fixed expenses relating to such items in an effort to boost consumers' willingness to pay. As firms' fixed costs increase because of outlays aimed at boosting product demands, however, their profits fall back, which removes the incentive for additional rivals to enter the market. In Sutton's framework with endogenous sunk fixed costs, therefore, steady increases in market size ultimately generate no further increases in the number of firms. Hence, concentration eventually reaches a lower bound as market size continues to expand.

Endogenous Sunk Costs and Concentration

An excellent exposition of the essential Sutton framework has been provided by Shiman (2007), who analyzes Sutton's theory in a setting in which market competitors engage in Cournot rivalry with respect to both their product quality choices and their levels of output. As in Sutton's (1991) original model, each firm in Shiman's specific framework decides whether to incur an exogenously fixed entry cost. In the first stage, a potential competitor decides whether or not to enter the industry in light of the entry cost that would be entailed. In the second stage, firms engage in Cournot rivalry concerning their quality choices, with a given firm that seeks to expand the demand for its product via a higher quality choice incurring a fixed cost of doing so. In the third stage, firms engage in Cournot rivalry regarding outputs of their products.

In the context of this setting with Cournot rivalry with respect to both quality and quantity, Shiman provides analytical solutions that generally correspond qualitatively to the specific simulated relationship displayed in Fig. 4.3. As market size initially expands along a relatively low range of values, there is an increase in the number of firms supported within a zero-profit equilibrium. Beyond a critical value of market size, however, the number of firms in the industry declines and eventually

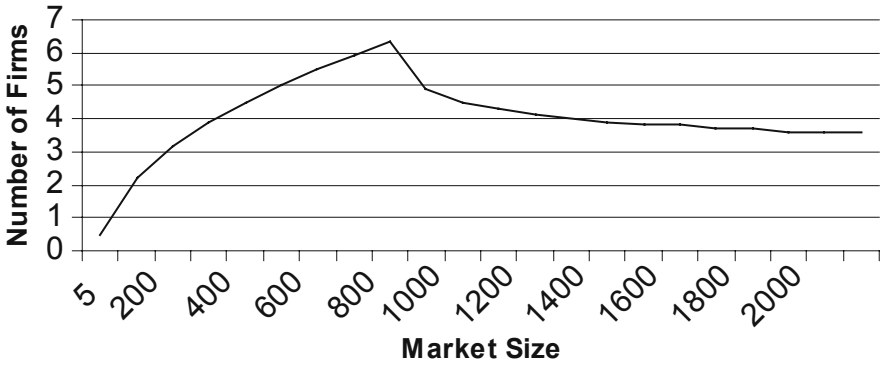

Fig. 4.3 The predicted relationship between the market size and the number of firms with endogenous sunk fixed costs
(Source: Simulation parameterization in Shiman (2007))

levels off at or below an upper bound. As suggested by Sutton, this implies a lower bound on concentration in an industry in which endogenous sunk outlays on product quality are an important dimension of market rivalry.

Non-Price Competition in Banking: Implicit Deposit Rates Versus Quality Rivalry

Sutton's theoretical framework applies to industries in which non-price competition is an essential feature. It has long been recognized in the banking literature that banks often compete along non-price dimensions as well as on the basis of explicit interest rates and service charges. In particular, when deposit rate ceilings were in place in the United States prior to rate deregulation in the early 1980s, numerous authors suggested that developing a framework of analysis of non-price competition was central to understanding rivalry in banking markets.

Initially, work along these lines focused on the idea that when confronted with legal deposit rate ceilings, banks offered implicit interest on deposits. For instance, Barro and Santomero (1972) measured the implicit deposit interest rate as a remission rate banks granted their customers. Benjamin Klein (1974) sought to estimate a deposit rate that banks counterfactually would have paid if not prohibited from doing so by legal constraints. Becker (1975) measured the implicit deposit rate as noninterest expenses less service charges as a percentage of deposits, and this became the standard measure utilized in most other analyses, such as Mitchell (1979), Startz (1983), Merris (1985), and Bradley and Jansen (1986).

As shown by VanHoose (1988), however, implicit deposit rates measured as remitted service charges arise in imperfectly competitive banking markets with or without deposit rate ceilings and must dissipate in perfectly competitive markets in which service charges fully adjust to reflect marginal costs of providing relevant services. Furthermore, as noted by Michael Klein (1978), such implicit rates are measured by economists ex post and cannot function as market signals ex ante, because depositors do not have detailed knowledge of bank costs to compute implicit rates of interest. Indeed, bank customers have no reason to care about bank costs per se but instead base their decisions on the characteristics of the products banks offer.

Thus, a more productive approach to contemplating how banks engage in nonprice competition is to focus on how banks seek to alter consumers' perceptions of the characteristics, or qualities, of their products. Heggestad and Mingo (1976), for instance, proposed that quality of service—such as weekly office hours of walk-in or drive-in services and availability of 24-hour automated services derived from survey data—influences behavior of bank customers and offered evidence suggesting an inverse relationship between bank concentration and service quality levels. White (1976) found a similar result using the number of bank branches as the relevant quality variable in U.S. banking data from 1970. Carlson and Mitchener (2006) suggest

that bank branches likewise were important elements of non-price competition in the 1920s and 1930s.

There is evidence that access to banking network in the form of branches or automated-teller-machine systems continues to play an important role in the competitive interplay among banks. Calem and Nakamura (1998) study U.S. deposit rate data from 1985 and from 1989 to 1990 and find evidence that although branch banking helps banks differentiate their product, it ultimately brings banks into more direct competition and thereby reduces localized market power. Hirtle (2007) focuses on branch network size and concludes that banks with mid-sized branch networks may be at a competitive disadvantage with respect to competitors possessing larger networks. In a study of bank-level data from Norway over the period 1988–1995, Kim and Vale (2001) conclude that the placement of branches is a crucial strategic variable for banks. They offer evidence that the relative size of a bank's branch network influences its market share but that such networks have meager feedback effects onto competitors and hence do not have an impact on overall market size. Cerasi et al. (2002) likewise conclude that branching was a key strategic variable in European banking markets during the 1990s, and Schmid (1994) found evidence in 1980s data from four European nations that banks branch to an extent dictated by consumer preferences. Dick (2006) explores the impacts of U.S. banking deregulation during the 1990s on branch banking. She finds that banks responded by competing through a significant expansion in the number of branches, which together with other adjustments in the deregulated environment boosted operating costs considerably. Increased branching generated greater revenues, however. On net, therefore, bank profits were unaffected and, additionally, there was virtually no adjustment in banking market concentration—a result consistent with Sutton's proposed endogenous-sunk-cost mechanism.

Evidence on Advertising Outlays in the Banking Industry

Sutton initially applied his theory to industries in which advertising is a commonly utilized approach to trying to boost consumers' willingness to pay. Could advertising outlays represent a form of endogenous sunk costs that make his theory applicable to the banking industry?

Unfortunately, there is a limited number of studies on the economic impacts of advertising in the banking industry. Martín-Oliver and Sals-Fumás (2008), for instance, examine whether advertising outlays by Spanish banks between 1983 and 2003 succeeded in boosting deposit supply and loan demand. They conclude that the answer is yes in both cases, although the magnitudes of responses are relatively small, with a deposit supply response with respect to advertising of 0.22 and a loan demand elasticity of only 0.11.

The other studies—Lapp (1976), Edwards (1973, 1976), Rhoades (1980), De Pinho (2000), DeYoung and Örs (2004), Hasan et al. (2002), Kohers and Simpson (1981), Örs (2006), Scott (1978) and Wolken and Derrick (1986)—have

focused mainly on the relationship between bank market concentration and advertising outlays. These studies reach contradictory conclusions. For instance, Rhoades and Kohers and Simpson find evidence of negative effects of concentration on advertising, but Hasan et al. and De Pinho conclude that there is a positive relationship. Edwards found no evidence of a relationship. Some studies, such as Scott's, find some evidence favoring an inverted-U-shaped relationship between concentration and advertising, implying greater advertising by banks in oligopolistic markets, while others, such as Wolken and Derrick, do not.

Several of these studies suffer from small-sample problems or reliance on survey data and consider banking data from much earlier periods. The Hasan et al. and De Pinho analyses apply to more recent data, and both conclude that there is a positive relationship between concentration and advertising expenditures at depository financial institutions, as do the more recent and broader studies by DeYoung and Örs (2004) and Örs (2006). DeYoung and Örs examine data from almost 1,900 U.S. thrift institutions in more than 600 deposit markets from 1994 through 2000 and find evidence of a positive association between market concentration and advertising outlays (except for mutual institutions). Örs evaluates data from nearly 4,500 commercial banks between 2001 and 2004 and additionally seeks to control for potential endogeneities between concentration and advertising expenses. He finds some evidence favoring an inverted-U-shaped relationship but a positive effect of concentration on advertising outlays within the sample. In addition, Örs concludes that advertising outlays per dollar of deposit decline as a bank's scale increases and that advertising has a positive effect on a bank's profitability.

Consequently, evidence from recent data covering a broader range of banking institutions indicates that advertising is an important facet of bank competition. This conclusion, in turn, provides another rationale—along with outlays for expansions of branches and ATM networks—that Sutton's theory of advertising outlays as a form of endogenous sunk costs arguably might be applicable to the banking industry.

Endogenous Sunk Costs and the Banking Industry

In light of the importance of branching as a strategic variable and evidence that advertising also is a competitive tool in banking markets, the banking industry appears to be a potential candidate "fit" for Sutton's model. Indeed, in a study examining the how fixed costs associated with establishing offices in banking markets help to explain variation in returns across markets, Hasan and Smith (1997, p. 48) speculate that "it is possible that existing banks use these fixed costs to deter entry and thereby raise profits." Furthermore, Dell'Ariccia (2001) builds on Dell'Ariccia et al.'s (1999) analysis of adverse-selection-based entry barriers to provide a theory of endogenous sunk fixed costs generated by adverse selection. Hence costs generated by adverse selections problems provide yet another possible explanation for concentrated structures of banking markets. In light of the abundance of motivations for endogenous sunk fixed costs in banking, Gual (1999) argues that bank market

structure ultimately depends on whether the main determinants of competition in banking are variable and exogenous versus endogenous sunk fixed costs.

Nevertheless, to date only the single study by Dick (2007) has directly examined the empirical relevance of Sutton's theory for the banking industry. Dick considers U.S. banking data encompassing more than 300 regional banking markets, with populations—her primary measure of market size—ranging from less than 100,000 people to more than 2 million. She finds little variation in bank market concentration across different markets sizes, as shown in Fig. 4.4, which is based on data provided in Table 2 of Dick (2007). As shown in panels (a) of the figure, the one-bank concentration (C1) ratio, measured as the fraction of deposits held by the largest firm in the regional market, varies relatively little around Dick's sample average of 0.30. Panel (b) shows that the same pattern holds true for the Herfindahl–Hirschman index (HHI) measure of bank market concentration, which exhibits little variation around the sample average value of about 1,850. Dick shows that scatterplots relating either

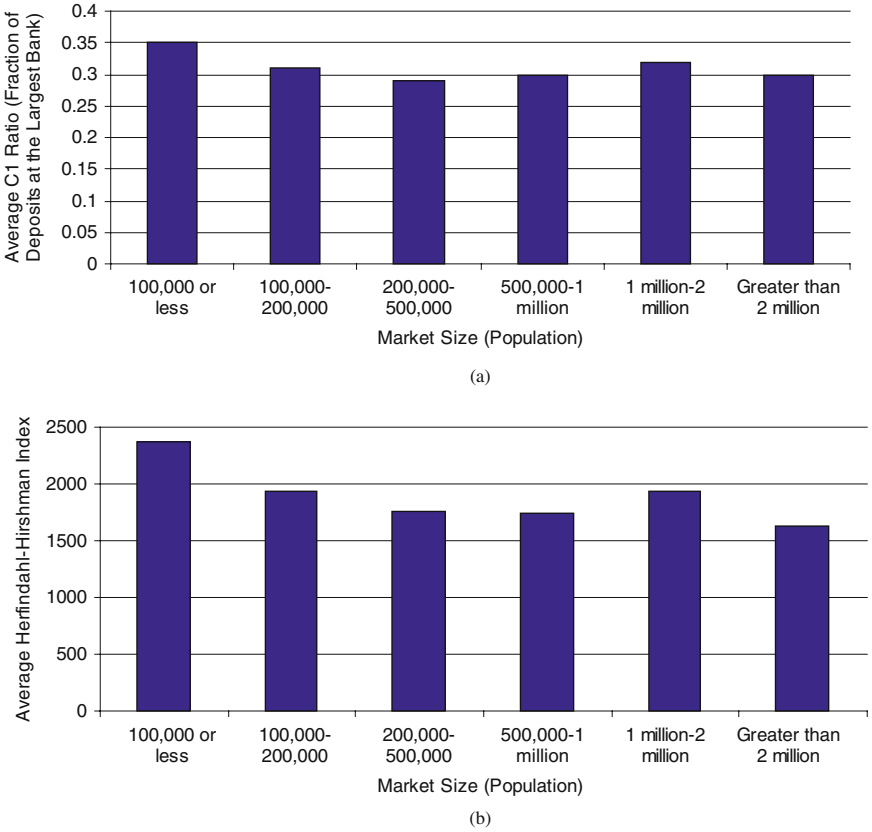

Fig. 4.4 The relationship between U.S. regional bank market concentration and market size (Source: Dick (2007))

the C1 or HHI to market size appear to imply a lower bound on market concentration. Estimated values of lower bounds on concentration depend on the presumed underlying distribution of markets. Nevertheless, Dick shows that under alternative distributional assumptions lower bounds are in fact implied by U.S. banking data.

Dick considers various measures of bank quality, including advertising intensity (outlays as a fraction of assets), branch density (branches per square mile in a regional market), and alternative measures such as employees per branch, salary per employee, and number of states in which a bank operates. She finds that, consistent with the Sutton model, each of these potential measures of quality increases with market size. She also concludes that within a given market, larger banks typically provide higher levels of quality than smaller banks do, a result consistent with an implication of Sutton's theory when extended to heterogeneous banks offering differentiated products.

Summary: The Industrial Organization of Banking

- The structure-conduct-performance (SCP) hypothesis follows directly from applying static oligopoly theories to analysis of bank market behavior. The SCP hypothesis suggests that barriers to entry that limit the number of loan- and deposit-market competitors engenders a rise in bank loan and deposit concentration and consequently generates imperfectly competitive conduct that results in higher market loan rates and lower market deposit rates, smaller loan and deposit quantities, decreases in consumer and producer surpluses, and increases in deadweight losses. Thus, the SCP hypothesis predicts that increased concentration induces worsened performance of the banking industry.
- A significant portion of the industrial-organization-of-banking literature has offered both cross-section and time-series evidence offering at least qualified support for the SCP predictions regarding loan and deposit quantities and interest rates.
- Consistent with the SCP contention that a product-differentiation barrier can restrain market competition, both past and recent studies have placed a spotlight on the bank–customer relationship. Early studies' predictions that forming long-lasting customer relationships allows banks to trade off lower loan rates and higher deposit rates in early periods for higher loan rates and lower deposit rates in later periods has received some limited empirical support.
- In recent years, researchers have focused more attention on whether banks can capitalize "soft" information gleaned from repeat customers in an effort to build market power. Some theories suggest that greater bank market competition eases the formation and continuation of customer relationships, while other theories make the opposite prediction. Empirical evidence on this issue is unsettled, although at present the weight of the evidence seems to support the conclusion that relationship lending has declined in the face of increased loan-market competition. Evidence regarding the importance of informational aspects

of bank–customer relationships is also mixed. Nevertheless, recent work suggests that informational effects of customer relationships may help to explain earlier studies' findings regarding impacts of distances of customers from banks.

- The efficient structure (ES) theory suggests that scale and scope economies could account for the existence of relatively large banking organizations. In contrast to the SCP hypothesis, the ES theory predicts that cost advantages of scale and scope accruing to such institutions could lead to lower loan rates and higher deposit rates even in relatively highly concentrated banking markets. Although evidence favoring significant economies of scale or scope in banking is mixed, a number of authors have offered empirical support for this fundamental prediction of the ES theory.
- Research by Sutton (1991) suggests that in industries in which endogenous sunk costs are empirically significant, a lower bound on concentration may exist. In the banking industry, possible sources of endogenous sunk fixed costs include commonly observed features of banking markets such as expensive branch networks, costly promotional advertising, and expenses incurred in addressing adverse selection problems. Dick (2007) provides evidence supporting the hypothesis of a lower bound on bank concentration consistent with Sutton's analysis.

Chapter 5
The Economics of Banking Antitrust

Antitrust policy involves a set of policies aimed at promoting competition in markets and thereby attaining allocative efficiency. Toward this end, antitrust policy traditionally encompasses efforts to prevent the formation and maintenance of price-fixing cartel agreements, to inhibit unilateral actions by any seller that would have the consequence of considerably enhancing its market power, and to avert mergers that would result in a significant lessening of competition and expansion of market power.

As discussed in the following chapter, banks typically operate under the watchful eyes of regulators, so under most circumstances, widespread banking cartels realistically could not function without governmental sanction. Actions by any individual institution to pursue formation of a monopoly also could be readily detected by regulators. Thus, the main focus of antitrust policy in banking—and hence of this chapter—is placed on efforts to forestall anticompetitive consolidations.

Why Banks Merge

Shull and Hanweck (2001) and Ingo Walter (2004) argue that the fundamental reason that banks contemplate and consummate mergers and acquisitions is to enhance shareholder value. In principle, consolidating institutions could be consistent with this objective if the merger boosts expected profits or, if shareholders are risk-averse, reducing risks.

Profit Enhancements from Mergers

Either an anticipation of a larger stream of revenues or an expectation of reduced average operating costs could lead shareholders to anticipate a greater discounted present value of profits following a banking consolidation. There are two fundamental mechanisms through which a bank merger could produce higher revenue flows for a consolidated institution than pre-merger institutions could obtain separately. One channel is more effective provision of income-generating services by the

D. Van Hoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-642-02821-2_5, © Springer-Verlag Berlin Heidelberg 2010

post-merger institution. If the consolidated bank is operated by better management than one or more of the consolidated institutions, then the consolidated bank should be able to expand its market share and increase its revenues.

Alternatively, a gain in market power can enable the post-merger bank to set interest rates and fees at levels that deviate further from perfectly competitive levels may have been the case prior to the merger. A fundamental source of market-power gains might arise if consolidation gives the post-merger institution pricing power not possessed by the pre-merger organizations. Other sources of market power gained from consolidation of resources might be more subtle. For instance, in the context of a Salop-style spatial-competition model, Hauswald and Marquez (2006) focus on customer relationships as a motivation for merger-and-acquisition activity in banking. Consistent with work by Sharpe (1990), they develop a theory suggesting that merger acquisitions enable banks to acquire proprietary information, which enables consolidated banks to moderate lending competition and expand market share, thereby increasing revenues and boosting profitability.

There are several ways in which a merger could reduce a bank's operating costs. An acquiring institution's replacement of an inefficient management might, instead of or in addition to improving managerial capability to generate additional revenues, bring about efficiency gains if new managers can improve the input mix and achieve X-efficiency gains by implementing lower-cost technologies and business methods. In addition, bringing together specializations of different banking firms under a single management structure could provide product-mix synergies that yield scope economies. Finally, consolidating into an absolutely larger organization could yield scale economies that reduce per-unit operating expenses.

Are the decisions to initiate bank mergers motivated more by predicted revenue enhancements or estimated cost savings? Houston et al. (2001) study large U.S. bank acquisitions between 1985 and 1996 and find evidence that the stock share prices of consolidated banks rose in relation to the pre-merger share prices of the pre-merger institutions, reflecting anticipated post-merger profitability increases. For 41 of these mergers, Houston et al. examine management projections of revenues and costs, and they conclude that post-merger share price increases were better explained by managerial estimates of cost savings than by anticipations of revenue enhancements.

Profitability motivations for bank mergers imply that acquired banks should share characteristics that identify them as likely takeover targets. Hannan and Rhoades (1987) examine a sample of acquired Texas banks between the early 1970s and early 1980s. Based on empirical analysis that employs a multinomial logit estimation procedure, Hannan and Rhoades conclude that larger market shares, lower capital-to-assets ratios, and location in urban markets are key factors raising the probability of a bank being an acquisition target. Likewise, Amel and Rhoades (1989) try to identify factors driving U.S. bank mergers between 1978 and 1983 and conclude that the lower a bank's earnings, the more likely it was to be acquired. Nevertheless, Palia's (1993) study of more than 130 U.S. bank mergers in the mid-1980s suggests that acquirers shy away from or pay lower premiums to obtain acquisitions with larger amounts of nonperforming loans. Palia concludes that key factors attracting attention to potential target banks are presence of the latter in highly concentrated markets—suggesting perceived rents generated by market power—and relative

difference in size between the acquiring and acquired banks—potentially implying a greater potential scope for investing in technologies that will provide significant revenue enhancements via broadening of the target bank's range of services.

Campa and Hernando (2006) study European banking consolidations between 1998 and 2002 and conclude that targets of acquisitions generally experienced lower pre-merger operating performance than average among their peers. Consistent with results of Cybo-Ottone and Murgia (2000) indicating positive abnormal returns of European bank merger announcements, Campa and Hernando find that return on equity increased by about 7 percent following on the heels of consolidations during the period of study.

Akhigbe et al. (2004) attempt to identify factors that made large, publicly traded banks within a sample of more than 250 acquisitions between 1987 and 2001 more desirable merger targets. They find a preponderance of acquisitions of larger banks with higher levels of core retail deposits but a lower return on assets and more non-performing loans, consistent with the interpretation that banks with considerable resources but relatively low performance levels are more likely to be acquisition targets. Inconsistent with most other studies, Akhigbe et al. find that better capitalized banks are more likely to be acquisition targets, a finding that Hannan and Pilloff (2006) suggest may result from Akhigbe et al.'s focus on large, publicly traded banks.

In contrast, in a study of U.S. bank consolidations between 1982 and 1999, Hadlock et al. (1999) fail to find that any particular governance incentive or performance variables are associated with a greater probability of a bank being acquired. They do find, however, that banks in which managers own greater portions of shares are less likely to be acquired, suggesting that entrenched managers may be able to block acquisitions that otherwise could be profitable to acquirers. Wheelock and Wilson (2000) examine U.S. bank acquisitions from the mid-1980s to the mid-1990s and, consistent with most other studies, find that banks with lower returns on assets are more likely to be acquired. They do not, however, find evidence that cost efficiencies are important factors driving acquisitions.

Although Cheng et al. (1989) also examine the characteristics of acquisition targets, which they find tend to perform below their peers, they focus attention on features of acquirers. They examine 135 U.S. banking takeovers during the first half of the 1980s and find that mergers in which acquirers pay higher prices in relation to the book value of assets involve larger, faster-growing acquirers with relatively high returns on assets. Acquirers with higher earnings ratios and measured market value-to-book, Cheng et al. conclude, also pay more for banks, suggesting that acquiring banks with better managers anticipate earning higher profits from acquisitions.

Hannan and Pilloff (2006) argue that the key determinant of whether or not bank merger acquisitions take place is the difference between the valuations of potential acquirers and target banks, which they argue offers acquirers a higher level of leverage that enables them to maximize post-merger performance gains relative to the costs of the consolidations. They examine data on more than 8,000 banking institutions, of which more than 1,400 were associated with acquisitions between 1996 and 2003. Like Cheng et al., Hannan and Pilloff find that larger institutions with greater market shares were more likely to be acquirers. Larger banks were also

more likely to acquire larger targets, and acquisition targets are more likely to have larger amounts of core deposits but relatively lower profitability. Consistent with their hypothesis regarding relative market values, Hannan and Pilloff also conclude that banks with higher capital-to-asset ratios are more likely to acquire banks that operate with lower capital-to-asset ratios. In a recent study of more than 150 bank mergers in the European Union between 1996 and 2004, Hernando et al. (2009) additionally find that acquired banks exhibit relatively weaker revenue performance and thereby add to the weight of evidence suggesting that lower-income banks are more likely to be targets in acquisitions. So does work by Pasiouras et al. (2007), who examine European Union banking mergers between 1997 and 2002.

Diversification Benefits of Bank Mergers

Another commonly cited factor providing an incentive for banks to consolidate their operations is potential gains in shareholder value generated by diversification of costs and risks. In theory, such gains in value may be attained both by broadening the scope of the consolidated bank's asset portfolio and by expanding the geographic scope of its operations. With regard to geographic diversification, mergers between banking institutions previously operating in separate markets may permit an acquiring bank to pool lending and funding risks. Diversification benefits are likely to be particularly significant if economic growth in a target institution's markets is negatively correlated with growth in the acquirer's markets, and work by researchers such as Rivard and Thomas (1997) has suggested that a geographic broadening of banks' activities is associated with lower volatility of earnings and reduced insolvency risk. Nevertheless, consolidating banking operations in pursuit of product and geographic diversity can potentially raise average costs at the merged institution if product-mix and geographic synergies fail to materialize, resulting in managerial and other operating inefficiencies, for which banks may be tempted to compensate by taking on greater risks in an effort to boost average returns.

To test whether mergers increase bank risk or yield risk-reducing benefits, Benston et al. (1995) examine data on 1981–1986 U.S. bank acquisitions involving acquirers with assets in excess of $100 million and acquisition targets with at least $25 million in assets. They conclude from an analysis of merger purchase premiums that the bank mergers they studied yielded risk-diversification benefits rather than expanding risk exposures. Hughes et al. (1999) conclude from an analysis of more than 400 U.S. bank holding companies that the benefit of risk diversification is the strongest argument in favor of bank consolidation.

Liang and Rhoades (1988) find in a study of more than 5,000 U.S. banking organizations between 1976 and 1985 that composite measures of risk—particularly financial risk, as predicted by portfolio theory—were reduced by increase geographic diversification. Liang and Rhoades also conclude that operating risks of banks increased in conjunction with broadened geographical diversification, as evidenced by lower earnings and lower capital-to-asset ratios.

In a merger-simulation study of risks at small U.S. community banks, however, Emmons et al. (2004) conclude that risk-diversification benefits of consolidation are likely to be dwarfed by risk reductions that would be brought about by greater scale of operations. In addition, there is little evidence that banking consolidations influence banks' exposures to interest-rate risks. Based on a study of 477 large U.S. bank mergers between 1980 and 1994, Esty et al. (1999) find no evidence of significantly different pre- and post-merger interest-rate-risk exposures.

Indeed, there is some evidence that higher-risk banks are more likely to engage in merger activities. In their study of banking mergers in the European Union between 1998 and 2002, Pasiouras et al. (2007) find that banks involved in mergers tended to be less well capitalized than banks that shied away from consolidation. Based on a study of about 1,000 German bank mergers during the 1995–2001 period, Koetter et al. (2007) conclude that most mergers involved banks with relatively high risk profiles, including a number of banks that officially were categorized as distressed institutions.

Furthermore, DeLong (2001, 2003) finds that U.S. bank mergers that diversify a consolidated bank's activities and spread them over larger territories typically do not raise shareholders' value. She finds instead that consolidations focusing on activities and geographical presence are more likely lead to enlarged earnings streams. Otherwise, she concludes that bank mergers ultimately add to shareholder value only when a merger involves either a relatively inefficient acquirer that experiences efficiency gains through the consolidation or partners that successfully reduce overall risk exposure and expected bankruptcy costs across the merged institution.

Assessing Loan and Deposit Market Effects of Bank Consolidation

From the point of view of banks' owners, mergers offer potential revenue enhancements, cost savings, and diversification benefits. Balanced against these benefits are potentially higher costs and risks that may be experienced by consolidated institutions once mergers have been consummated.

What are the broader implications of banking mergers for society as a whole? How do bank consolidations affect market outcomes and impinge on the welfare of banks' borrowers and depositors? Let's first contemplate the answers that basic economic theory offers to these questions.

Mergers in Initially Perfectly Competitive Banking Markets

Theoretically, bank mergers in initially competitive markets can result either in net social welfare deteriorations or improvements. To see why this is so, see Fig. 5.1. The figure simplifies by assuming that banks' marginal resource costs are identical and invariant to banks' balance-sheet choices. Thus, at the outset in the loan market,

in panel (a) banks face the total marginal and average unit cost of lending equal to $r_N + C_N^1 + C_L^1$, which in a perfectly competitive loan market corresponds to the market supply of loans. The initial equilibrium quantity of loans therefore is L_1, and the market loan rate is r_L^1. At the outset in the deposit market, as shown in panel (b), the per-unit net marginal and average return on deposits for each identical bank is $(1 - q)(r_S - C_S^1) - C_D^1$, which corresponds to the market demand for deposits by banks. The initial equilibrium quantity of deposits therefore is D_1, and the market loan rate is r_D^1.

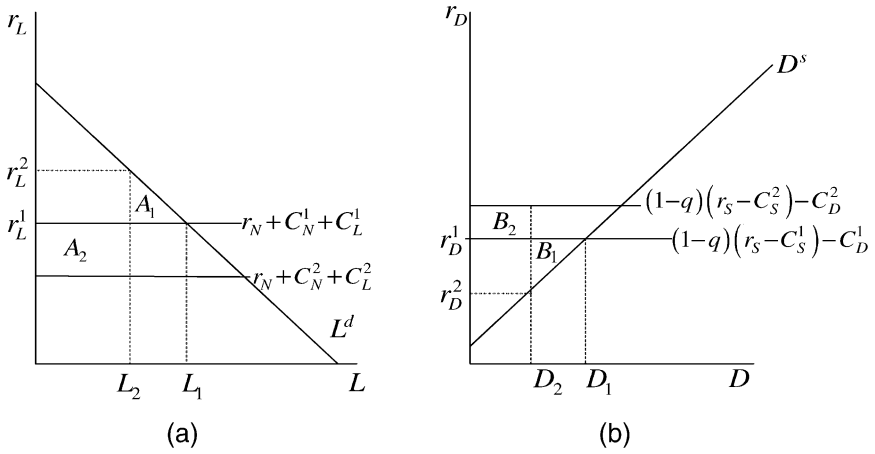

Fig. 5.1 Effects of industry consolidation in initially perfectly competitive loan and deposit markets

Now suppose that there is consolidation among banks through mergers. Such consolidation may have two sets of effects, both of which are depicted in Fig. 5.1. One is an increase in market power. As shown in panel (a), consistent with the structure-conduct-performance (SCP) paradigm discussed in the previous chapter, greater market power results in a reduction in lending, to L_2, a higher market loan rate, r_L^2, and consequently, a decrease in consumer surplus equal to the triangular area denoted A_1. The outcomes in the deposit market, as displayed in panel (b), are a decrease in deposits, to D_2, a decline in the deposit rate, to r_D^2, and a reduction in producer surplus equal to the triangular area labeled B_1.

A second set of effects of industry consolidation could, as suggested by the efficient-structure theory, be lower-cost production of financial services. If so, after consolidation the banks' marginal resource costs will decline to smaller values denoted C_S^2,C_L^2,C_D^2, and C_N^2 in both panels of Fig. 5.1. In panel (a), the result is a loan resource cost savings equal to the rectangular area A_2, and in panel (b) the outcome is a boost in net revenues derived from deposits equal to the rectangular area B_2.

Banking industry consolidation generates a decrease in total surplus in the loan market if A_1 exceeds A_2, but loan-market surplus can rise if the reverse is true. Likewise, total surplus declines in the deposit market if B_1 is greater than B_2

but rises otherwise. Thus, in principle, sufficient improvements in bank operating efficiency—that is, sufficiently large reductions in C_S^2,C_L^2,C_D^2 , andC_N^2 generated by banking industry consolidation of resources—can bring about gains in surplus that exceed the reductions in surplus resulting from greater market power of merged institutions. If so, on net such industry consolidation can be socially beneficial.

Mergers in Initially Imperfectly Competitive Banking Markets

Now let's suppose that banking industry consolidation occurs via bank mergers in initially imperfectly competitive loan and deposit markets. Figure 5.2 considers the implications of industry consolidation in this alternative setting. Thus, market power initially possessed by banks results in a market loan rate, r_L^1, above the total marginal and average cost of lending and, consequently, a level of lending, L_1, beneath the level the level that would have prevailed under perfect competition. In the deposit market, market power possessed by rivals at the outset yields a market deposit rate, r_D^1, below the net marginal and average revenue derived from deposits, which results in a deposit level, D_1, that is lower than the quantity of deposits that would have been observed under perfect competition.

Again, consolidation among banks via mergers generates two sets of effects. There is an increase in market power, so panel (a) depicts a lending reduction, to L_2, a rise in the market loan rate, to r_L^2, and a decrease in the sum of consumer and producer surplus equal to the trapezoidal area denoted A_1. In the deposit market, panel (b), shows that merger consolidation brings about a reduction in deposits, to

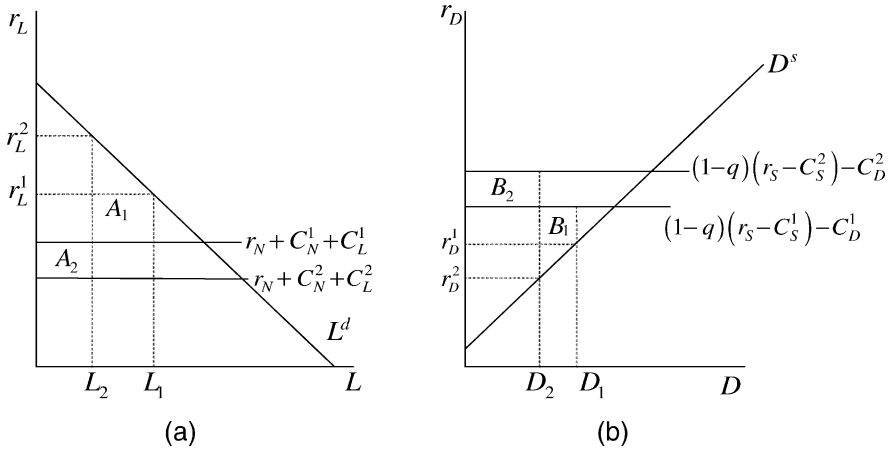

Fig. 5.2 Effects of industry consolidation in initially imperfectly competitive loan and deposit markets

D_2, a decrease in the deposit rate, to r_D^2, and a decrease in the sum of producer surplus and net revenues from deposits to banks equal to the trapezoidal area labeled B_1.

A second set of effects of industry consolidation arises if consolidation permits more efficient production of financial services, in which case the banks' marginal resource costs fall to the values labeled C_S^2,C_L^2,C_D^2, and C_N^2 in both panels of Fig. 5.2. In panel (a), the result is a loan resource cost savings equal to the rectangular area A_2, and in panel (b) there is a rise in net revenues derived from deposits equal to the rectangular area B_2.

Once more, banking industry consolidation brings about a reduction in total surplus in the loan market if A_1 exceeds A_2, but loan-market surplus rises otherwise, and total surplus falls in the deposit market if B_1 is greater than B_2 but increases otherwise. Again, cost efficiencies achieved through industry consolidation can bring about gains in surplus that exceed the reductions in surplus resulting from greater market power of merged institutions.

Note, however, that the areas A_1 and B_1 in Fig. 5.2 are now trapezoidal rather than the correspondingly labeled triangular areas in Fig. 5.1. It is more likely in this case, therefore, that for similar changes in interest rates and marginal resource costs, reductions in total surplus will result from banking consolidation. Thus, efficiency gains achieved through industry consolidation are less likely to yield gains in social welfare if the industry is imperfectly competitive prior to consolidation via mergers than if the industry is perfectly competitive at the outset.

Evidence on the Consequences of Banking Consolidation

The pre- and post-consolidation welfare comparisons conducted in Figs. 5.1 and 5.2 suggest that the net market impacts of bank mergers are ambiguous on theoretical grounds. Thus, assessing these net impacts ultimately is an empirical issue.

Mergers and Market Power

Does banking consolidation contribute to increased market power? Most research seeking to address this question has focused on the U.S. experience. In a study of U.S. banking mergers and acquisitions in the 1980s and 1990s, Berger et al. (1999) conclude that consolidation in the U.S. financial services industry is indeed associated with greater market power. They find evidence of increased profit efficiency and better risk diversification as a consequence of bank mergers but no evidence of significant cost efficiency gains. Nevertheless, Berger et al. (1998) conclude that mergers and acquisitions during this period did not necessarily reduce lending to small businesses once reactions of incumbents are taken into account along with refocusing efforts of the merged institutions.

Garmaise and Moskowitz (2006) examine U.S. data involving 316 bank mergers and acquisitions during the 1992–1999 interval and conclude that the results of

these consolidations were higher loan rates and less lending that ultimately contributed to decreases in real estate values and even higher crime in affected market regions. Scott and Dunkelberg (2003) study the credit market experiences of small U.S. business borrowers during the mid-1990s, of whom about 25 percent borrowed from banks that were involved in mergers during that period. They conclude that mergers neither substantively affected the ability of firms to obtain credit nor significantly altered the loan rates they confronted. Nevertheless, Scott and Dunkelberg find evidence that following mergers, small business borrowers faced more stringent nonprice loan terms, dropoffs in service quality, and more or higher fees for services. On the deposit-market side of the consolidation issue, Simons and Stavins (1998) utilize survey data on deposit rates and find that mergers generate lower deposit rates. Interestingly, the find that the most pronounced effect of mergers is on other rival banks, which appear to respond to the reduction in the number of market rivals brought about by mergers by reducing rates paid on deposits. In an examination of monthly U.S. banking data encompassing mergers between 1997 and 2006, Craig and Dinger (2007) find evidence of downward pressure on rates paid on transactions deposits. Park and Pennacchi (2008) examine U.S. data on acquisitions of small banks by large multimarket banks between 1994 and 2005 and likewise conclude that these consolidations reduced deposit rates.

In principle, mergers also may impact the availability of banking services through branch networks. Avery et al. (1999) find evidence that there is a per capita reduction in U.S. bank branches within a ZIP-code area containing overlapping branches of merged institutions, suggesting some reduction in banking services following a merger. They do not, however, find an overall negative relationship between banking consolidation and the per capita number of banking offices. Becher and Campbell (2005) note that a number of bank mergers during the 1990s ultimately failed to create shareholder value, particularly in situations involving a high degree of branch overlap that experienced significant negative returns.

Sapienza (2002) examines effects of Italian bank consolidations on credit outcomes between 1989 and 1995. She finds that borrowers of small banks acquired in a merger tended to benefit from better loan terms. In general, she concludes that loan interest rates initially decreased as a consequence of bank efficiency gains but ultimately increased over time as the merged banks exercised their increased market power, particularly over borrowers with only a handful of additional banking relationships. Di Patti and Gobbi (2007) evaluate more than 450 Italian bank mergers during the 1990s and find that mergers generated declines in lending, especially when they resulted in termination of customer relationships at acquired banks.

There is considerable evidence that impacts of banking consolidation on borrowers and depositors differ depending on whether they are customers of an acquiring bank or an acquired bank. Carow et al. (2006), for instance, analyze data on more than 2,000 corporate borrowers from banks involved in the ten largest U.S. merger combinations during the 1991–2001. Based on stock price reactions of loan customers of acquiring and acquired banks, they conclude that customers of acquired banks experienced negative outcomes. In an analysis of the effects of mergers and acquisitions involving more than 200 Italian banks between 1984 and 1996,

Focarelli et al. (2002) find that mergers are driven by strategies aimed at broadening services with an aim to increase revenues and hence profits, while acquisitions push down bad loans at acquired banks and thereby boost the acquired banks' profitability. Focarelli et al. conclude that in both types of consolidations, lending to small businesses declined. In a study of Norwegian bank mergers between 1983 and 2000, Karceski et al. (2005) consider the effects of merger announcements of on the share prices of the banks' borrowers. They find evidence of abnormal positive returns for borrowers of acquiring banks but abnormal negative returns for borrowers of acquired banks, which suggest that the mergers' strategic motivations favored acquiring banks.

In addition, Karceski et al. find evidence that they argue is consistent with borrowers with low switching costs responding to mergers by terminating relationships with acquired banks. This finding is consistent with work by Sharpe (1997), who constructs a theoretical model aimed at predicting the fraction of customers that switch their patronage among firms under different market structures. Sharpe tests these predictions using data from a panel of more than 200 U.S. banks in about 100 markets during the mid-1980s. He concludes that for this sample, increased deposit market concentration was associated with lower deposit rates and that this effect was enhanced when a larger portion of depositors failed to shift allegiances to other banks as a consequence of switching costs, which Kiser (2002a, b) finds, based on consumer surveys of 1,500 U.S. households, to be greatest for households on either side of educational and income means.

How large are the costs of switching faced by customers of acquired banks? Kim et al. (2003) utilize a panel of Norwegian bank data during the period spanning 1988–1996 to estimate borrower demand and bank lending relationships implied by a theoretical model of oligopolistic banking rivalry. They estimate that during this period in Norway, borrowers faced an average percentage switching cost of about 4.1 percent. This amount was about one-third of the average market loan rate for the sample, which implies significant scope existed for an acquiring bank to push up loan rates without losing customers.

Evidence on Efficiency Gains from Banking Consolidation

Is the enhancement of market power at least partly counterbalanced by efficiency gains from bank mergers? In a study of all U.S. banking consolidations involving at least $1 billion of assets for both partners during the 1980s, Akhavein et al. (1997) conclude the mergers boosted managerial efficiency with respect to earnings. Merged banks, they found, experienced an average increase in profit efficiency of at least 16 percent relative to other large banks, with most profit-efficiency improvements generated by revenue enhancements resulting from shifting more assets from securities to loans rather than from changes in interest rates or fees. Cornett et al. (2006) also focus on profit efficiency effects of banking consolidations. They examine a number of U.S. bank consolidations between 1990 and 2000 and conclude that operating profit performance gains were greater for mergers that involved larger

rather than smaller banks, that focused on a single activity instead of diversification of activities, and that were more intensively geographically focused rather than geographically broadening. Cornett et al. find evidence of both revenue enhancements and cost reductions, but the former tended to be more pronounced than the latter.

Other authors suggest that significant cost savings are often realized in banking consolidations. Humphrey and Vale (2004) apply a flexible cost function to 1987–1998 data on 130 Norwegian bank mergers and find that on average these mergers generated cost savings. In an analysis of the U.S. banking consolidation trend during the 1980s and 1990s, documented by Rhoades (2000a, b), Berger et al. (2007) conclude that cost efficiency gains were particularly significant for large banks. In a study of cost and profit efficiencies of more than 7,000 U.S. banks between 1993 and 1998, Berger and DeYoung (2001) find only modest efficiency gains from geographical expansion through merger and other means, however.

There is some evidence that in contrast to the conclusions of Berger et al. (1999), Garmaise and Moskowitz (2006), Scott and Dunkelberg (2003), Park and Pennacchi (2008), Sapienza (2002), and Di Patti and Gobbi (2007), efficiency gains from banking consolidations can more than offset market power effects. Erel (2007) examines proprietary data on commercial and industrial loans at 300 U.S. banks and 50 branches of foreign banks involved in mergers between 1990 and 2000. He finds that on average, mergers reduced loan spreads and that the reduction was larger at acquiring banks with greater declines in post-merger operating costs. In addition, Focarelli and Panetta (2003) analyze the post-merger behavior of deposit rates in Italy through most of the 1990s and find that although short-run deposit rates dropped in the short run, in the long run efficiency gains dominated over market power impacts, which resulted in higher deposit rates. Ashton and Pham (2007) study 61 bank mergers in the United Kingdom between 1988 and 2004 and conclude that there were substantial cost efficiency gains for the consolidated institutions without significant effects on most retail interest rates.

Evanoff and Örs (2009) suggest that bank mergers and acquisitions can generate efficiency gains at non-merging incumbent banks, which respond to news of a merger by improving the efficiency of their operations in an effort to remain viable competitors. To test this hypothesis, Evanoff and Örs develop performance measures for all U.S. commercial banks between 1984 and 1999, and they examine how measured productive efficiency of non-merging banks responded to consolidations of other banks during this period. They find support for their hypothesis, particularly among incumbent banks facing acquisitions of rival banks by non-local institutions for which the acquisitions constitute entry into the affected markets. Efficiency improvements at non-merging incumbents were also significant in cases in which banks previously possessed the greatest market power, indicating that incumbents had been protected from external competition prior to an acquisition.

Banking Antitrust in Practice

Since 1963 and 1964, when the U.S. Supreme Court ruled in the cases involving the Philadelphia National Bank and the First National Bank and Trust of Louisville, bank mergers have been subject to two key antitrust laws—the Sherman Act of 1890, which first forbade efforts to monopolize a market, and the Clayton Act of 1914, deemed unlawful specific business activities, such as certain forms of price discrimination and exclusive dealing. The Bank Holding Company Act of 1956 already prohibited the Federal Reserve's Board of Governors from "approving a proposal that would result in a monopoly" or that would "substantially lessen competition in any relevant market" subject to exceptions in situations in which "the probable effect of the proposal" meets the "needs of the community to be served." Taken together, the 1963 and 1964 Supreme Court decisions and the 1956 legislation placed the Federal Reserve at the center of bank antitrust policies. The Bank Merger Acts of 1960 and 1966 broadened oversight to include the other two key U.S. banking regulators, the Office of the Comptroller of the Currency (OCC) that supervises nationally chartered banks and the Federal Deposit Insurance Corporation (FDIC) that regulates state-chartered banks that are not members of the Federal Reserve System. Nevertheless, these other agencies have typically followed the Federal Reserve's lead in the antitrust policy sphere.

U.S. Bank Merger Guidelines

Under terms of the Bank Merger Act of 1966, any contemplated banking merger must first be proposed to the applicable banking regulator (or regulators in cases in which the planned merger involves banks supervised by different regulatory agencies). If the applicable regulator renders a favorable judgment, then the merger cannot be completed for another 30 days, pending potential further review by the U.S. Department of Justice and private parties possessing legal standing to challenge a merger under U.S. antitrust laws. If no legal challenges arise by the end of this 30-day waiting period, then parties to the merger can engage in the proposed consolidation. At any point, however, banking regulators or the Department of Justice can act to avert the proposed consolidation.

The Relevant Market

According to procedures developed by the Federal Reserve's Board of Governors (Walter and Wescott, 2008; American Bar Association, 2007) each of the twelve Federal Reserve banks defines the relevant banking markets—clearly delineated, distinct markets for banking services—to be utilized in evaluating effects of any proposed mergers within the geographic area encompassed by its own Federal Reserve district boundaries. In keeping with the U.S. Supreme Court's 1963 decision emphasizing the importance of local market considerations, in defining the relevant banking market staff economists and other officials at Federal Reserve

banks emphasize geography in their delineations. Thus, as an initial estimate they typically rely considerably on the U.S. Office of Management and Budget's so-called "MSAs"— Micropolitan Statistical Areas containing at least one urban area and a population between 10,000 and 49,999 people and Metropolitan Statistical Areas containing at least one urban area and a population of 50,000 or more people—and on Rand McNally's Ranally Metropolitan Areas, or "RMAs," containing at least 70 people per square mile with at least 20 percent of its labor force commuting to a defined central urban location. In some cases, the Federal Reserve takes into account political county boundaries as well.

Once a geographic region has been identified as the relevant banking market, the relevant item sold in that market by banks must be identified. The 1963 Supreme Court ruling determined that item to be "the cluster of products and services" offered by banks but did not specify a definition of the appropriate "cluster" to be considered in banking antitrust analysis. In practice, banking regulators and the Department of Justice have chosen to utilize bank deposits as a surrogate measure of the appropriate cluster of banking services. The Federal Reserve analyzes concentration of the relevant market utilizing data on deposits that banks report as of June 30 each year. It excludes deposits at banking institutions specializing solely in credit card lending, because these institutions raise their deposit funds in nationwide markets.

To assess the degree of concentration in the relevant banking market, the Federal Reserve follows the Department of Justice by relying on the Herfindahl–Hirschman index (HHI), the sum of squared percentage market shares of each of the banks in the market. It computes both pre-merger and post-merger HHIs, and it also calculates the resulting change in the HHI, which in the case of a merger of two rivals turns out to equal twice the product of percentage market shares of the two rivals. (The pre-merger HHI is equal to $\sum_{i=1}^{n} (S_i)^2$, where there are n banking institutions in the relevant market and S_i is the market share of bank i. The post-merger HHI in the event of a merger by the first two banks, $i = 1, 2$, must equal $\sum_{i=1}^{n} (S_i)^2 - (S_1)^2 - (S_2)^2 + (S_1 + S_2)^2 = \sum_{i=1}^{n} (S_i)^2 + 2S_1S_2$; thus, subtracting the pre-merger HHI, $\sum_{i=1}^{n} (S_i)^2$, yields $2S_1S_2$.) Both the post-merger HHI and the change in the HHI are important elements in banking regulators' assessment of a proposed merger.

Merger Screening

In evaluating planned consolidations of nonfinancial firms, the U.S. Department of Justice considers a market to be unconcentrated if its HHI is less than 1,000, to be moderately concentrated if its HHI is between 1,000 and 1,800, and to be highly concentrated if its HHI exceeds 1,800. Figure 5.3 suggests that relying on these Department of Justice classifications of market concentration would imply that many U.S. banking markets are already relatively concentrated. As the figure indicates, the average HHI values for urban banking markets (defined for simplicity as MSAs) and rural banking markets (defined as non-MSA counties) have been

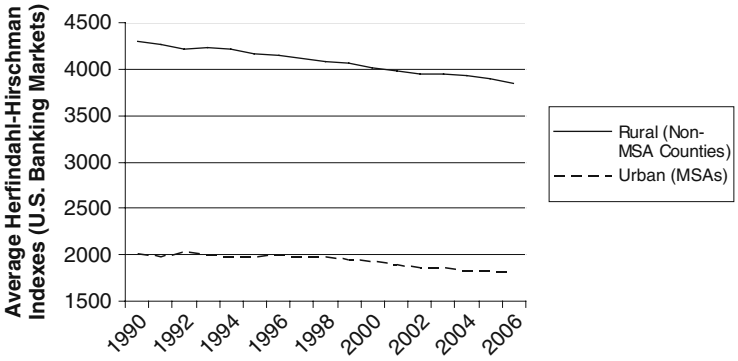

Fig. 5.3 Average Herfindahl–Hirshman index values in U.S. banking markets (Source: Pilloff (2009))

trending downward recently. Nevertheless, average HHI values in urban markets have remained very close to the Department of Justice's minimum level of 1,800 for defining high market concentration, and average HHI values for rural markets are well above this value.

Indeed, most U.S. banking mergers would fail to meet the Department of Justice's screen for assessing proposed nonfinancial mergers, according to which a planned merger raises antitrust concerns if post-merger HHI level exceeds 1,800 and the change in the HHI resulting from the merger exceeds 50. Proposed U.S. banking mergers, however, face a different screening procedure than do planned mergers of nonfinancial firms.

The U.S. bank merger screening mechanism consists of two screens. The first, known as Screen A, calculates the HHI for the relevant banking market giving 100 percent weight to the deposits of commercial banks but placing only a 50 percent weight on deposits of savings institutions, which U.S. banking authorities perceive to provide a narrower cluster of banking products than commercial banks. This naturally reduces the effective value of the post-merger HHI that a proposed bank merger would yield. Under screen A, a proposed banking merger would raise antitrust concerns on the part of banking regulators and the Department of Justice only if the resulting post-merger *weighted* HHI exceeds 1,800. In addition, in recognition of the fact that in the course of their day-to-day operations banking institutions face some competition from savings institutions and other myriad financial institutions, antitrust concerns about a proposed merger arises if the change in this weighted HHI exceeds a value of 200.

If the Screen A thresholds for the post-merger weighted HHI and change in weighted HHI are exceeded, then the Department of Justice applies a second screen, called Screen B. This screen computes a post-merger *unweighted* HHI and change in the *unweighted* HHI for RMAs instead of the relevant markets as defined by the Federal Reserve. If the thresholds are again exceeded under Screen B, then antitrust policymakers consider other factors specific to the planned merger in reaching their final judgment regarding its merits. Among these so called "mitigating

factors" are whether there currently is direct competition between the merging institutions, whether the merging institutions appear to specialize in providing different types of services, whether customer surveys suggest that the merger partners' clusters of products are close substitutes, and whether the locations of their offices and customers exhibit significant geographic overlap within the relevant market.

As a matter of procedure, within the Federal Reserve System, Board of Governors can authorize a Federal Reserve bank to approve a merger application, but only the Board of Governors has authority to seek to block a proposed consolidation. As noted above, however, even if the Federal Reserve initially approves a planned merger, the Department of Justice and private parties with legal standing can seek to block it in court under U.S. antitrust laws.

Evaluating the U.S. Bank Merger Guidelines

The formal merger guidelines utilized by banking regulators and the U.S. Department of Justice provide a relatively clear roadmap for banks contemplating a consolidation. Indeed, a Federal Reserves Web site, http://cassidi.stlouisfed.org/, provides access to HHI data that prospective partners can utilize to assess the prospects for approval of a contemplated merger.

Nevertheless, the bank merger guidelines also raise several economic questions. Let's begin with the first and perhaps most important of these, which is the guidelines' definition of the relevant market.

Is the Official Relevant Banking Market Really Relevant?

It has long been understood that a central issue of antitrust analysis is determination of the relevant market. Stigler (1955, p. 4) argued that from an economic perspective, an industry's market

> should embrace the maximum geographical area and the maximum variety of productive activities in which there is a strong long-run substitution. If buyers can shift on a large scale from product or area B to A, then the two should be combined. If producers can shift on a large scale from B to A, again they should be combined.
>
> Economists usually state this in an alternative form: All products or enterprises with large long-run cross-elasticities of either supply or demand should be combined into a single industry.

From an economic point of view, therefore, a proper delineation of the relevant market for antitrust analysis should encompass both product space and geographical space and should take into account substitution capabilities with respect to both consumption and production.

A Theory-Policy Mis-Match

As noted by Stigler, significant, positive long-run price elasticities of demand and supply should provide evidence of substitutability. In considering whether products

X and *Y* purchased and sold within a common geographical area are close substitutes in consumption, the finding of a positive long-run cross-price elasticity of demand at current prices would indicate that a rise in the price of product *Y* would induce consumers to purchase more of item *X*. A positive long-run price elasticity of supply between goods *X* and *Y* at current prices would imply that an increase in the price of good *X* would firms producing good *Y* to switch some of their resources to production of good *X*. Presumably, this product-substitutability test could be expanded to encompass the issue of geographical space by computing cross-price elasticities of demand and supply for ever-larger geographic areas in which items *X* and *Y* are bought and sold, to the point at which the magnitudes of the elasticities become economically insignificant.

In the merger guidelines that it applies to nonfinancial firms, however, the Department of Justice defines a market to be the minimum group of products and smallest geographical area such that a hypothetical monopoly of all related products in the area could raise price by 5 percent above within a one-year period. As Stigler and Sherwin (1985) note, the Department of Justice guidelines

> . . .are inconsistent in defining markets. For instance, producers who could use existing facilities to enter within six months are included in the market even though their competitive influence may be less than that exerted by similar goods that would not have been included in the market because the 5 percent test was already satisfied. Similarly, current geographic sales patterns will be used to make the initial selection of the geographic market even though areas will be included that would not have been under the 5 percent test. Further, and paradoxically, the Guidelines' 5 percent test will ensure that markets with prices currently above the competitive level are defined more broadly than otherwise identical markets experiencing competitive pricing.

Moreover, Stigler and Sherwin argue, the Department of Justice methodology is inconsistent with a price-data-oriented, economic approach to defining markets.

Geographical Arbitrariness in the Bank Merger Guidelines

Application of the Department of Justice guidelines to banking markets is fraught with additional difficulties. First, no monopoly-pricing-increase criterion—5 percent or otherwise—is entailed in applying the guidelines to defining banking markets. Instead, only relatively arbitrary geographic-space considerations are taken into account in adapting the guidelines to the banking industry, consistent with Kwast et al.'s (1997) contention that large fractions of households and small businesses within a given local area tend to cluster purchases of key financial services at banking institutions also located within the same narrow area. Relating to this point, in their review of recent research on the relevant banking market for antitrust analysis, Gilbert and Zaretsky (2003) conclude that most evidence supports the conclusion that distance is indeed a key factor in determining the scope of interactions among banks in their offerings of product clusters to consumers, implying that it is reasonable to view local communities as the relevant market for most banks. Consistent with this judgment, Brevoort and Hannan (2006) examine bank lending patterns in nine U.S. metropolitan areas between 1997 and 2001. Brevoort and

Hannan conclude that distance is a key lending deterrent, particularly for smaller banking institutions, and that the importance of distance may have increased in recent years.

Nevertheless, Gilbert and Zeretsky (2003) also suggest that evidence points to a potential broadening of relevant markets, both geographically and in terms of the range of competitors that banks face. Indeed, Jackson (1992) examines deposit market definitions founded on metropolitan statistical areas using U.S. data from the mid-1980s and concludes that while interest rates on retail transaction and savings deposits are locally determined, rates on six-month certificates of deposit are determined in a national market. Jackson argues that this conclusion suggests that assuming that the local market is relevant for all banking services results in a market definition that is overly narrow. Based on data from the mid-1990s, Radecki (1998) argues that banking markets typically extend to state levels rather than being narrowly confined to cities or counties. In addition, Heitfield and Prager (2004) examine intrastate differences in deposit rates by U.S. banks in 1988, 1992, 1996, and 1999. Although Heitfield and Prager conclude that local market concentration is the main determinant of bank deposits rates, they also find evidence that state-level concentration plays an important role. Edelstein and Morgan (2006) argue that a better indicator of banking market dimensions than the deposit rate may be the amounts that banks pay to buy one another's branches, and they find that branch sale prices in ten northeastern states have been more closely correlated with concentration at the state level rather than the local level. Amel and Starr-McCluer (2002) also note that household survey evidence indicates that between the late 1980s and late 1990s, there was a noticeable upward drift in distance between households and the financial institutions they frequent, particularly for nondepository services—that is, a wide range of productive activities that involve bank outputs rather than inputs.

These results suggesting a broadening of the geographic scope of U.S. banking markets are consistent with implications of recent work by Berger and DeYoung (2006). They examine U.S. multibank holding companies between the mid-1980s and late 1990s and conclude that there were significant reductions in agency costs per unit of distance from headquarters offices to affiliate locations and branches. This agency-cost decline, Berger and DeYoung suggest, resulted in more centralized control over affiliates. In principle, this closer direction of affiliate and branch networks by managers in otherwise dispersed home offices could produce greater overlap in strategic rivalry—and hence, effective competition—among banking organizations that extends beyond localized regions that the guidelines presume to be relevant markets.

Do the Formal Guidelines Mis-Measure Market Power?

The U.S. bank merger guidelines utilized by the Federal Reserve and other banking authorities and by the Department of Justice presume that a reasonable proxy for market power is market concentration as measured by the HHI. As discussed in Chapter 4, this presumption is consistent with the structure-conduct-performance paradigm, but it is supported neither by evidence that various researchers have found

to be consistent with the efficient structure theory. In addition, use of concentration as a proxy for market power is inconsistent with modern approaches to measuring market power, such as those reviewed by Perloff et al. (2007). Furthermore, in their study applying one of these modern approaches—conjectural variations—to panel data from more than 450 Norwegian banks during the early 1990s, Berg and Kim (1998) argue that relying on concentration measures such as the HHI can provide unreliable indications of market power. Berg and Kim's analysis points to strategic interdependence in retail loan markets but relatively competitive behavior in the corporate loan market. Nevertheless, they find that relative HHI values are reversed from the relative concentration levels that their findings would have suggested.

Even within the context of the SCP paradigm, some researchers have questioned whether the HHI is the appropriate measure of the degree of competition in banking markets. Hannan (1997), for example, seeks to evaluate whether the Herfindahl–Hirschman index (HHI) accounts sufficiently for separate effects of market-share variation and number of rivals in explaining observed loan and deposit rates. He finds no evidence that the HHI is inadequate in analysis of market deposit rates but concludes that market loan rates are better explained by the number of competitors than by the HHI. More recently, Carbó-Valverde et al. (2009) consider a cross section of banking data from more than 1,900 banks in 14 European countries between 1995 and 2001. They conclude that five measures of bank market competition—the net interest margin, the Lerner index (the difference between product price and marginal cost as a proportion of the price), return on assets, the H-statistic (a measure of the extent to which input cost changes are incorporated into price changes), and the HHI—are only weakly positively related. By implication, the HHI is only weakly related to variables for which the formal bank merger guidelines assume that it is reasonable proxy.

Implications of Endogenous Sunk Fixed Costs

As discussed in Chapter 4, Dick (2007) provides evidence that endogenous sunk costs could be an important characteristic of competition in the banking industry. If this conclusion is justified, then banking effectively could be viewed as a sort of "natural oligopoly," and antitrust analysis may be overly focused on industry concentration. In particular, the current critical post-merger HHI value of 1,800 and change in HHI of 200 specified by U.S. bank merger guidelines conceivably could be inconsistent with the banking industry's "natural" structure—a conclusion reinforced by separate evidence of a slowing change in the size distribution of U.S. banks documented by Janicki and Prescott (2006). Furthermore, in principle the average level of quality of banks' products could be enhanced by mergers, implying that more concentrated banking markets conceivably could yield higher welfare for consumers.

Based on the results of her study, Dick (2007) suggests that an important implication for U.S. antitrust policy with respect to banking is that bank quality is a key variable that should be incorporated into policy evaluations of the consumer welfare impacts of changes in structure within banking markets. According to Dick,

"if consumers are receiving higher quality and are benefiting as a result, they are not necessarily hurt by. . .higher prices they have to pay" for the higher quality. By implication, if a bank merger that results in a more concentrated market also boosts the fixed costs associated with the provision of better products made possible, both *prices* and consumer welfare could rise. In principle, therefore, violations of the critical HHI thresholds specified by U.S. bank merger guidelines actually could yield higher-quality banking products and improvements in consumer welfare.

Much depends, Dick suggests, on whether a proposed merger will maintain rivalry among a few dominant banks and potentially several fringe competitors. Her analysis indicates that as long as scope for active quality competition remains, consumer welfare will not necessarily be harmed by a merger that creates a more concentrated banking market.

Do Banking Consolidations Preclude Entry and Reduce Consumer Welfare?

Recent work in industrial organization by Perloff et al. (2007) and others has suggested that dynamic game-theoretic considerations should be crucial elements in assessing the market-power and consumer-welfare implications of mergers. So far, the banking literature has yet to incorporate a number of these latest developments.

There have, nonetheless, been a few important recent contributions that have progressed in extending several advances in industrial organization to the realm of banking markets. One area that recent work has explored regards whether banking markets are open to entry of new rivals. If so, the SCP paradigm's traditionally static perspective on the effects of market consolidation perhaps should carry less weight in antitrust policy.

On the surface, the evidence on bank entry appears to favor the static view implicit in the SCP paradigm. Earlier work by Rhoades (1997) has pointed out how lack of customer information, the presence of customer switching costs, and sunk costs and other frictions associated with entry can constitute significant barriers to bank entry. Rhoades discusses strategic entry barriers that banks may erect, including exclusive contracts for minibranches in retail chain stores and branch site preemption. In one of two recent studies of the entry issue, Berger and Dick (2007) examine study data on 10,000 U.S. bank entries into local markets between 1972 and 2002. They find that on average there is an economically significant early-mover entry advantage accruing to U.S. banks, which they estimate to be as high as 15 percentage points of market share. In the other study, Adams and Amel (2007) broaden the definition of entry beyond creation of a new banking institution to include the opening of branches in new markets by banks established in other markets. In spite of this expansion in the definition of "banking market entrants," Adams and Amel find, based on a consideration of U.S. banking data between 1994 and 2006, that considerable changes in market conditions would be required to generate even slight increases in the probability of entry, particularly in rural markets.

Does variation in bank market structure have measurably significant effects on consumer welfare? Recent work by Dick (2008) casts light on this issue. Under simplifying assumptions about consumer preferences, Dick utilizes logit techniques

developed by Berry (1994) and Berry et al. (1995) to estimate consumer welfare in more than 300 U.S. banking markets, defined therefore in some accordance to the formal antitrust guidelines as metropolitan statistical areas, during the period of considerable structural changes in banking spanning the middle and latter 1990s. She finds that changes in welfare over this interval varied relatively little across more or less concentrated markets. Furthermore, she concludes that impacts on consumer welfare of service enhancements such as expansions of branching networks were at least as significant as changes in welfare caused by variations in interest rates and fees. This result suggests that banking antitrust policy's traditional focus on the latter variables might overlook the importance of other determinants of consumer welfare also affected by banking consolidations.

Rethinking Bank Merger Analysis

Ultimately, there is no general resolution to the contrasting predictions of the structure-conduct-performance (SCP) and efficient-structure (ES) theories regarding net empirical relationships among concentration, competition, and welfare. In some instances, SCP-predicted effects on quantities, interest rates, and welfare in banking markets may emerge as a result, say, of merger-and-acquisition activity that boosts industry concentration and thereby generates reduced competition. In other situations, ES-predicted effects on quantities, interest rates, and welfare may emerge through enhanced cost efficiency that produces relatively larger firms and hence more observed concentration without necessarily decreasing the intensity of market competition. Consistent with these mixed theoretical predictions, the empirical literature offers considerable evidence that mergers can enhance market power but also indicates efficiency gains can occur. These mixed findings regarding the impacts of mergers on bank interest rates, fees, and quantity choices indicate that the net welfare effects of mergers likely vary on a case-by-case basis.

In a review of Whinston (2006), Hall (2007) notes that modern antitrust law

> is gradually shifting toward the principle that an antitrust case is a demonstration that customers would be better off without the merger than with it. Modern courts are losing their single-minded devotion to the formulaic approach of defining a relevant market, measuring market power within that market, and only then considering the effects of conduct challenged as harmful to competition. In place of that rigid formula, modern courts would like to know by how much the conduct has raised prices or diminished product quality.

Hall concludes that

> [a]s a practical matter, sponsors of a merger gain more traction at the [FTC and Justice Department] from a direct demonstration of a favorable or neutral effect on prices than they do by defining a relevant market and measuring the change in concentration in that market, following the recipe in the Guidelines. This is visible in Whinston's discussion, where the analysis needed to apply the market-definition principles overlaps substantially with the analysis needed to measure the unilateral effects of a merger. Soon, the Guidelines will read, "The FTC and Justice Department review proposed mergers by estimating the effects of the merger on the prices and other characteristics of all products affected by the merger."

As Hall notes, the two main contending methods for estimating merger effects on a case-by-case basis are merger simulation and event studies assessing effects of announcements of proposed mergers on share prices of the prospective merger partners, their rivals, and their customers. Merger simulation offers the potential to assess mergers in advance, based on pre-merger data (for a review of merger simulation modeling, see Budzinski and Ruhmer, 2008). Because implementation of the U.S. Department of Justice's merger guidelines focuses on likely post-merger price changes, most merger simulation models in the industrial organization literature, some of which have been applied in actual antitrust cases, assume Bertrand–Nash-style price-setting behavior (see Peters, 2003). Walker (2005) points out that post-merger pricing outcomes predicted by merger simulations based on Bertrand–Nash behavior potentially are very sensitive to variations in own and cross-price elasticities of demand. Thus, slight errors in consumer demand estimations can result in wide swings in simulated post-merger outcomes.

Whinston (2007) observes that in principle a market consolidation could cause market prices to move outside of the range of prior data, necessitating relying on out-of-sample extrapolations that can make demand estimations utilized for simulating merger outcomes unreliable. Whinston also calls attention to the fact that post-merger dynamics could deviate from those that prevailed prior to the consolidation; for instance, pre-merger strategic behavior along Bertrand-Nash might switch to Cournot–Nash behavior following a merger, in contrast to a merger simulation's assumption that Bertrand–Nash behavior would persist. Furthermore, Budzinski (2008) argues that the comparative fits to data of proposed merger simulation models that yield mutually incompatible merger predictions can be the same, implying that at best merger simulation models might be utilized to complement current techniques used by litigants in merger cases. Whinston (2007) concludes, nonetheless, that "[i]t seems clear that as techniques for estimating structural models get better, merger simulation will become an increasingly important tool in the analysis of horizontal mergers."

In the case of the banking industry, Molnár (2008) examines price-cost margins forthcoming from alternative assumptions regarding strategic behavior in banking markets in Finland between 2003 and 2006 and determines that market power in the Finnish banking industry mostly accords with Bertrand–Nash behavior assumed in most prior merger simulation models. Molnár proceeds to utilize the estimated Bertrand–Nash framework to conduct merger simulations and analyze the resulting effects on loan and deposit interest margins. Under the assumption that hypothetical merger pairings among the largest three banks in Finland would yield no cost efficiencies passed through to loan and deposit rates, he concludes that the average effects would be a 70-basis-point increase in the market loan rate (which translates into a 39 percent rise in the loan-rate margin) and a 10-basis-point decrease in the market loan rate (which implies a 37 percent increase in the deposit-rate margin). Molnár notes that these estimated impacts are relatively small in light of the significant market concentration—about 70 percent—that would result from such mergers and the maintained assumption of no cost efficiency benefits forthcoming from the simulated consolidations.

In contrast to Molnár's simulation exercises, cost-efficiency effects are central to the analysis of McIntosh (2002). In his study, McIntosh uses time-series revenue and cost data from Canada's five largest banks to estimate scale-efficiency effects and impacts on a banking-services price index that would have resulted if proposed mergers involving two pairs of these banks had been approved rather than denied by the nation's Competition Bureau and minister of finance in 1998. In contrast to most of the merger simulation literature, McIntosh assumes Cournot–Nash behavior, and under this assumption his simulation analysis indicates that the proposed mergers would have yielded cost efficiencies sufficient to mitigate increased market power in Canada's banking markets, resulting in lower consumer prices.

There have been a number of event studies of bank mergers. Several of these are reviewed by Rhoades (1994), who concludes that the potential for short-term abnormal stock-price movements to reflect speculation aimed at maximizing near-term trading gains rather than market perceptions of longer-term impacts of mergers. Whinston (2006) also notes that an alteration of the industry environment that helps motivate a merger yet has nothing to do with competitive considerations could also benefit rivals, implying the potential for false signals of diminished competition if stock prices rise for both merger partners and rivals. As Hall observes, however, stock-price events involving economically and statistically significant empirical results can provide strong evidence regarding a merger's likely impacts.

Antitrust Issues in Bank Payment Networks

Radecki (1999) examines the 1996 income statements of the 25 largest U.S. bank holding companies and estimates the revenues generated the banks' payments-related activities. Revenues generated by funds transfers on behalf of depositors, Radecki finds, accounted for 28 percent of the banks' total operating revenues. Credit-card payment processing generated 7 percent of their revenues, and securities payment processing contributed an additional 3 percent. All told, therefore, Radecki concludes that transactions relating to payments processing contribute generated about 38 percent of the banks' revenues.

In light of banks' traditionally significant role as payment intermediaries, Lacker and Weinberg (2003) and Kahn and Roberds (2009) review several studies related to banks' role in payment systems and suggest that more work should be done on "payment economics"—albeit from a central banking perspective with a slant on monetary theory rather than industrial organization. Some studies relating to payment issues that matter most acutely to central banks, such as VanHoose (2000) and Holthausen and Rochet (2006), examine central bank payment-system issues from an industrial organization point of view. Nevertheless, and perhaps unsurprisingly, most payment-economics work related to the monetary economics literature focuses on funds-market issues more germane to monetary theory and policy. This orientation perhaps helps to explain why some monetary economists have even applied

overlapping-generations models to the study of payment flows in overnight funds markets.

There is, however, a substantial and rapidly expanding body of industrial organization research on the economics of two-sided markets, including card payment networks that include a substantial number of banks as active participants. Not surprisingly, this industrial-organization-oriented analysis of payment economics is much more relevant from the perspective of banking antitrust than the monetary literature on the topic.

Bank Cards and Two-Sided Markets

Chakravorti (2003) provides a helpful review of analyses of theories of the economic functioning of card payment networks. Most of the theories that he reviews, and virtually all subsequent work, are based on the concept of two-sided markets. So let us begin our discussion of antitrust issues relating to payment networks by explaining how theories of two-sided markets are applied to networks that specialize in handling payments initiated using payment cards commonly issued by banks.

Card Payment Networks as Two-Sided Markets

To consider the essential economics of a two-sided card payment market, consider Fig. 5.4. In a two-sided market, one or more *platforms* facilitate dealings between *end users*. In the case of a card-payment network, the network itself—for instance, Visa, MasterCard, Discover, or American Express—constitutes the platform. The end users are retailers and individual cardholders.

Figure 5.4 depicts an environment in which a card network potentially charges end users both usage fees (u_c and u_r) and membership fees (m_c and m_r). Let's ignore the latter for the time being (see a discussion of membership fees below) by supposing that membership fees are equal to zero, so that the card payment network collects only usage fees. It does so by collecting from a cardholder an additional per-dollar fee u_c for each dollar of payments that it transmits to a retailer on the cardholder's behalf. Out of each dollar of payments transmitted to retailers, however, the card network deducts an amount equal to u_r and thereby charges this usage fee to retailers. Thus, the total per-unit usage fee charged to end users is $u = u_c + u_r$.

As discussed by Rochet and Tirole (2006a), If the aggregate volume of transactions processed by the card network platform were to depend solely on this total price—that is, if a change in u_c with u kept unchanged leaves aggregate card transactions unaffected—then the market would only be one-sided. If the aggregate volume of transactions varies with u_c while u remains unchanged, however, the market is two-sided: A change in the relative price structure alters total activity on the network.

Network Usage Externalities

In the bank payment card industry, the values that end users assign to access to a given payment network depends on the number of other end users. For instance, an individual who owns a card entitling her to access payment-clearing services of a network benefits when others opt to use the same card, because this increases the likelihood that retailers will choose to accept the network's card. Thus, economists say that products such as payment services provided via card payment networks are subject to network externalities, or third-party spillover effects onto some network participants resulting from decisions by others with whom they engage in no direct transactions

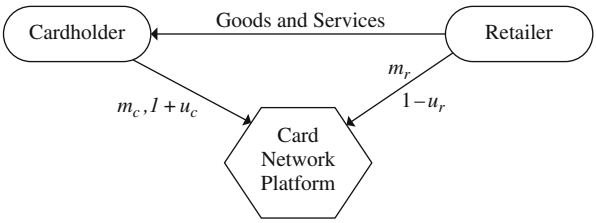

Fig. 5.4 A two-sided card payment network

Why do retailers accept payment cards? Undoubtedly, the primary explanation is transactional benefits—in the form of sales that otherwise would not take place—accruing to the retailer. Hayashi (2006), for instance, motivates retailers' acceptance of credit and debit cards by the upward shifts in product demands confronting individual merchants, which thereby yield incremental sales. In an analysis of retailers' motivations in accepting credit card payments, Chakravorti and To (2007) propose an additional incentive. They emphasize the revolving function of credit cards and propose that competing merchants accept credit cards because doing so permits them to close current sales with illiquid consumers rather than confronting uncertainty regarding future sales. Through recognition of actual and anticipated sales expansions, Rochet and Tirole (2006b) suggest, retailers essentially internalize the benefits experienced by customers as a generating benefits to themselves as well and hence opt to accept a payment network's cards as long as these and other potential benefits exceed the network's fees. Of course, since the cardholder benefits that retailers seek to internalize are subject to network externalities, merchant end users also confront externalities in card payment systems.

The two-sidedness of card payment networks' pricing structures hinges on the existence of these external spillovers. As discussed by Rochet and Tirole (2006a), the Coase theorem indicates that in the presence of clearly defined property rights and in absence of transaction costs and asymmetric information, private parties should be able to establish contracts that internalize usage externalities and attain

a social optimum. Two-sidedness of the price structure of card payment networks, Rochet and Tirole suggest, results primarily from transaction costs and constraints established by the networks or policymakers that prevent implementation of such welfare-improving contractual agreements.

Membership Externalities and Membership Fees

The act of providing a card payment network entails fixed expenses unrelated to variable usage of the network. Retailers incur fixed costs to establish card-acceptance systems, and acquirers experience significant fixed expenses in developing systems for clearing payments. Issuers also incur considerable fixed costs related to marketing and advertising. All of these fixed costs influence decisions regarding whether or not to participate in a card payment network, thereby determining the magnitudes of external benefits experienced by end users.

It is well known from the literature on price discriminating two-part tariffs and from the Ramsey pricing literature that including access fees can efficiently enable the recouping of fixed costs. This fact undoubtedly helps to explain why membership fees such as those depicted in Fig. 5.4 are often observed within the overall pricing structure of card payment networks. Rochet and Tirole (2006a) also point out that if a desired pricing structure includes a role for cross-subsidization, access prices such as membership fees can be employed to capture portions of end-user surpluses for such purposes. Armstrong (2006) provides a recent model in which flat membership charges emerge as important tools for balancing the externalities faced by end users in two-sided markets.

Multiple Actors and Pricing Structures in Card Payment Networks

As explained by Schmalensee (2002), the manner in which usage fee structures of a card payment network are established hinges on whether the network is a proprietary network—that is, owned and operated as an individual profit-maximizing "closed network," as in the case of the American Express card network—or a cooperative network—that is, owned and operated by an association of coordinating institutions, such as the Visa bank card "open network." The basic two-sided structure depicted in Fig. 5.4 is most applicable to a *unitary* proprietary network that engages in all issuing and acquiring activity on its own behalf. As Schmalensee notes, an alternative proprietary system is a *non-unitary* system in which the network contracts with other parties, such as a selected set of institutions to do some of the issuing and/or acquiring. In contrast, in a cooperative card payment network, independent institutions establish a contractual arrangement for coordinating card payments via the network.

Figure 5.5 provides a stylized depiction of co-existing card-payment networks. The figure displays two institutions, Bank *i* and Bank *j*, that are members of two networks, Card Network Platform 1 and Card Network Platform 2. For the sake of

simplicity, the figure shows Bank *i* as an issuer with respect to both card network platforms and displays Bank *j* as a payment acquirer with respect to the two platforms. In fact, some banks perform both roles in cooperative networks, although increasingly banks that operate as acquirers contract with third parties to process payments.

Note that pricing-structure complexity is greater in Fig. 5.5. The banks continue to charge membership and usage fees cardholders and retailers. In addition, the networks charge two types of fees to member banks. The fees b_1 and b_2 denote membership and/or usage fees that the network charged the banks, and the fees f_1 and f_2 denote interbank fees—commonly called interchange fees that the network collects on behalf of an acquiring bank from an issuing bank. Baxter (1983) provided the first formal analysis of interchange fees within a bank payment-clearing association. Baxter showed that the socially optimal interchange fee must take into account the two-sided nature of the payment transactions. In Baxter's model with perfectly competitive net-issuing and net-acquiring banks, the association itself is indifferent regarding alternative levels of the interchange fee—a result sometimes called "interchange-fee neutrality" in the literature on card payment networks.

Pricing Structures within a Non-Competing Card Payment Network

As discussed by Schmalensee and by Rochet and Tirole (2002), allowing for imperfect competition results in considerable scope for variation in pricing structures across card payment networks. Based on the analysis of card payment systems

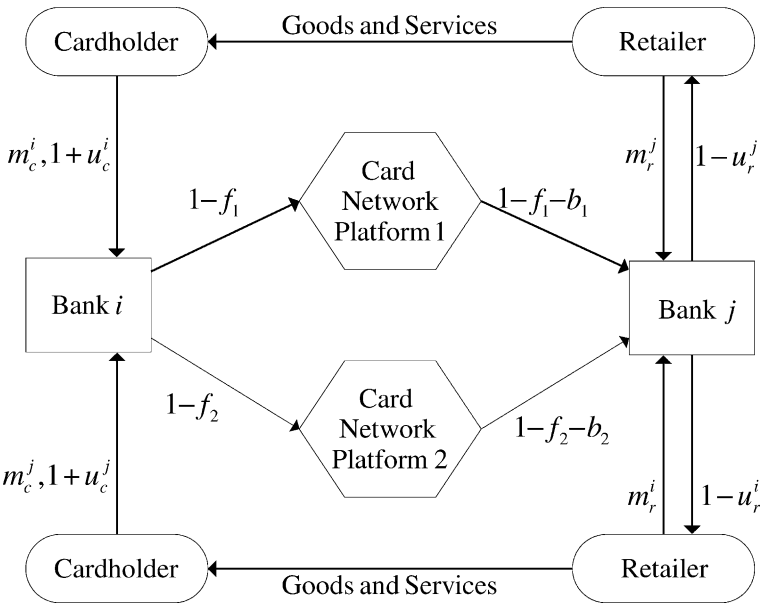

Fig. 5.5 Competing card payment networks

provided by Evans and Schmalensee (1999), most studies of networks' two-sided pricing structures presume low search costs and little product differentiation on the acquiring side of card payment networks, which researchers often offer as support for assuming perfectly competitive behavior among acquiring institutions. (As Evans and Schmalensee document, however, concentration among acquirers has increased noticeably.) Search costs and the scope for product differentiation is presumed to be higher on the issuing side, resulting in some market power for issuers

To minimize complexities, some studies, such as those of Rochet and Tirole (2002), Bolt and Tieman (2006), and Wright (2003, 2004) assume that rival card payment networks process a volume of transactions that is either arbitrarily fixed or essentially predetermined via an exogenously specified distribution of consumer transaction choices. Rochet and Tirole (2002) further examine a setting in which issuers have no incentive to operate as acquirers, which naturally further simplifies their framework of analysis. In the context of their model, Rochet and Tirole find that in the absence of other pricing complications such as surcharges for purchases with cash instead of cards, the interchange fee that maximizes overall social welfare is never higher than the profit-maximizing level. Indeed, without surcharges the profit- and output-maximizing interchange fees coincide, implying that if anything outcomes can arise in which there is a socially inefficient overprovision of payment cards. When card surcharges are permitted within the Rochet–Tirole framework, there is an underprovision of payment cards, and welfare implications are ambiguous.

Wright (2003) allows for heterogeneity in retailers' benefits from network participation not considered by Rochet and Tirole and concludes that attaining the welfare-maximizing interchange fee must balance differing externalities faced by heterogeneous cardholders and retailers. As a consequence, the profit-maximizing aggregate usage level typically lies below the socially optimal level. Wright (2004) also concludes that the welfare effects of imposing a "one-price policy" in which surcharges for card purchases are prohibited depends on the amount of market power possessed by retailers. If retailers possess market power, then the ability to impose surcharges would enable retailers to set prices to extract surplus from cardholders, thereby reducing revenues accruing to issuers. In the absence of retailer market power, however, a one-price policy results in a bifurcation of the retailing market into cash-only and card-only segments.

Schmalensee (2002) also studies the functioning of a single card payment platform in the context of a basic usage-pricing framework. He finds that a single unitary proprietary system a unitary proprietary platform that simply maximizes its own private value establishes a set of fees that generates the smallest transaction volume. Essentially, a unitary proprietary platform acts as a profit-maximizing monopoly and restrains output via a structure of sufficiently high fees that restrains production of transaction services to maximize its profits. In contrast, in Schmalensee's model, a cooperative system that aims to attain an output objective while covering operating costs balances benefits from spreading payment volumes across separate net-issuing and net-acquiring groups utilizing the payment network.

Within this single-network framework that abstracts from banks' interactions with cardholders and presumes fixed weights in network decision-making by banks that are either issuers or acquirers, Schmalensee derives a decomposition of the optimal interchange fee into two parts, with one part depending on the difference between demand elasticities across cardholders and retailers and the other part depending on the difference in costs faced by issuers and acquirers. He concludes that a cooperative platform typically will set a pricing structure that yields, relative to unitary and non-unitary proprietary platforms, the highest volume of payment transactions. A double-marginalization problem confronts multiple participants in both a cooperative network and a non-unitary network, with the latter producing a volume of payment transactions that is intermediate between the volumes forthcoming from cooperative and unitary proprietary networks.

Bolt and Tieman (2006) study a model of a monopoly two-sided platform such as the one depicted in Fig. 5.4 and hence ignore the issue of interchange fees. In addition, they abstract from membership fees. In this simplified setting, they demonstrate that when the socially optimal pricing structure is implemented, the monopoly platform fails to cover its operating costs. Thus, they conclude, network externalities faced by the card payment platform create a second-best pricing problem analogous to that faced by a natural monopoly, in which first-best, allocatively efficient pricing generates negative economic profits. A possible consequence, Bolt and Tieman suggest, is a card payment network pricing structure that includes cross-subsidization schemes, higher interchange fees, and no-surcharge rules preventing retailer participants from giving discounters to customers who pay with cash.

Competing Card Payment Networks

When card network platforms compete for end users, such as the case of a card-network duopoly illustrated in Fig. 5.5, the range of complexities faced in analyzing pricing structures broadens considerably. Nearly all of the work analyzing such competitive-network settings within the context of two-sided-market frameworks has appeared very recently. In these settings, end users can engage in multihoming: Cardholders can participate in more than one network by utilizing more than one payment card, or retailers can accept cards from more than one network from their customers.

Rochet and Tirole (2002) briefly touch on the impacts of competition between payment systems via multihoming by cardholders—that is, holding multiple credit cards issued by different networks. They contend that if one network seeks to undercut another but setting a slightly lower interchange fee, then retailers have an incentive to respond by accepting the former system's card and rejecting the card of the other system, potentially reducing welfare relative to the monopoly case.

In a more general two-sided-market framework, Rochet and Tirole (2003) examine a situation in which "buyers" (cardholders in the payment-network context) utilize the services of differentiated platforms that either are profit-maximizing firms or non-profit associations. Their model applies to an extension of the setting in Fig. 5.4 to the case of more than one platform competing for the same end

users. Rochet and Tirole include buyers and sellers of different types, with a network platform competing to obtain "marquee buyers"—cardholders that generate a particularly high surplus to retailers—from another platform and taking into account the existence of "captive buyers" that remain loyal to their initial platform choice. Other things being equal, the presence of marquee buyers naturally raises the seller price, but the presence of capital buyers skews the pricing structure to the benefit of sellers. Naturally, an increase in predisposition to multihoming on the part of buyers also is beneficial to sellers. These results could be sensitive, however, to Rochet and Tirole's assumption of no fixed membership-type fees, which Armstrong (2006) suggests could complicate joint determination of pricing structures in competing two-sided markets. Rochet and Tirole find that under some circumstances interchange fees could be lower and total market surplus greater in a monopoly setting than when two networks compete.

Cabral (2005) suggests, however, that Rochet and Tirole's ambiguous welfare results regarding monopoly versus competition between two-sided networks does not necessarily rule out the existence of other more-than-offsetting welfare losses that might arise from existence of a monopoly network. Furthermore, a recent analysis by Chakravorti and Roson (2006) studies a variety of interactions between payment networks, including duopoly and cartel behavior with or without symmetry of the networks' payment instruments (such as credit cards offered by one platform and debit cards on another). In the context of their framework, which includes some narrowing assumptions such as monopolistic retailers, they conclude that competition unambiguously raises the welfare of end users, although not generally uniformly, even in the case of symmetric competition.

Guthrie and Wright (2007) attempt to provide a classification of the wide array of equilibrium pricing-structure outcomes that can result from competition between two card identical payment platforms depending on the natures of potentially multihoming cardholders and retailers and on their behavior. In general, Guthrie and Wright find that either both networks set the same structure of fees with at least one side of end users (cardholders or retailers) multihoming, or only one network survives and attracts the exclusive participation of all end users. Guthrie and Wright conclude that in a setting with homogeneous retailers, competition between networks cannot boost fees charged retailers. They argue that if, in contrast, retailers are heterogeneous, platform competition can result in higher fees being assessed on retailers, thus leading to higher interchange fees. The fee structure can be further biased against retailers if consumers knowing that retailers will accept more than one card in equilibrium choose a preferred card to hold and thereby benefit from platform competition that focuses on attracting cardholders rather than retailers. In contrast, if retailers know that cardholders always hold multiple cards in equilibrium, retailers can "steer" cardholders to their preferred network, resulting in platform competition focusing on attracting retailers, to the benefit of the latter group of end users.

Regulatory and Antitrust Issues in Card Payment Networks

In recent years, card payment networks have been subjected to a variety of lawsuits, legal regulations, and antitrust actions. On the legal and regulatory front most of the attention has been placed on interchange fees. In 1979, National Bancard Association unsuccessfully challenged the existence of interchange fees in U.S. courts. A few years ago, the Reserve Bank of Australia banned no-surcharge rules and began regulating interchange fees. In Denmark, interchange fees on domestic card payments are prohibited, in Sweden and the Netherlands, no-surcharge rules are illegal, and in the European Union the European Commission recently intervened to bring down interchange fees. On the antitrust front, U.S. payment card networks have faced antitrust suits by the Department of Justice challenging restrictions on issuers and a proposed merger between acquirers. They also failed to prevail in lawsuits challenging the legality of "Honor All Cards" rules they previously had imposed requiring retailers accepting a network's credit cards to accept its debit cards as well (see, for instance, Economides, 2008).

There has been some initial work evaluating such regulatory and antitrust issues. In Schmalensee (2002), the levels of interchange fees under either cooperative or non-unitary proprietary systems are driven primarily by different weights within the network on the part of net-issuing versus net-acquiring banks. In the context of his analysis of a single card network platform, Schmalensee finds that the proprietary, private-value maximizing interchange fee theoretically may be higher or lower than an output-maximizing that a cooperative network platform might set. Furthermore, his analysis suggests that a policy action that might raise the interchange fee conceivably could boost network transaction output and that a policy action that reduces the interchange fee potentially could depress network transactions output. Policy intrusions into the fee-setting process, he concludes, could place cooperative networks at a competitive disadvantage with respect to proprietary networks, leading institutions to abandon higher-output cooperative networks in favor of lower-output proprietary networks. Schmalensee's conclusion, therefore, is that there is no clear economic argument in favor of antitrust policy that interferes in the setting of interchange fees.

As noted by Rochet and Tirole (2008), regulation of interchange fees is often motivated as requiring cost-based fees, but as they point out, governments have failed to apply the same logic to television, newspaper, videogame, and other two-sided industries. Furthermore, Rochet and Tirole (2006b) question whether retailers actually experience harm as a result of interchange fees and suggest that some perceived harm may result from the fact that heterogeneous retailers internalize customer benefits at different levels.

Schwartz and Vincent (2006) utilize a card-network model with variable transaction volumes to examine the implications of no-surcharge rules. They argue that such rules imbalance the fee structure between end users, resulting in harm to retailers, with networks responding by reducing cardholder fees and granting rebates to card users if possible, with cardholder welfare potentially declining if rebates are

not feasible. Network profits increase, and overall welfare increases only if there a sufficiently large number of cash users.

Rochet and Tirole (2006c) examine the effects of an honor-all-cards rule. In their benchmark model based on Rochet and Tirole (2002), they find that in fact such a rule benefits not only the multi-card network platform that imposes it but also allows the platform to optimally re-balance externalities across end users. The result is greater volumes of both credit and debit payments processed by the network, resulting in greater social welfare. Generalizing their framework to allow for heterogeneous retailers, differentiated platforms, and varying substitutability between credit and debit transactions yields the same re-balancing effect on the mutli-card network's pricing structure. The implications for welfare, however, tend to become ambiguous depending on parameter values.

Emch and Thompson (2006) contemplate the application of the Department of Justice merger guidelines to card payment networks. Emch and Thompson show that it is possible that monopolization of a two-sided market considerably raises the prices charged to cardholders but not retailers, or *vice versa*. They derive price markup formulas for fees charged to end users. Based on these relationships, they propose applying the guidelines' test for a small but significant and non-transitory increase in price to the sum of the prices the network charges to both groups of end users. White (2006) notes, however, that the pricing relationships derived by Emch and Thompson hinge their assumption of fixed proportions. He also concludes that their results likely depend on the simple two-sided framework they utilize, which more readily applicable proprietary networks than to open card associations.

Sun and Tse (2007) apply a differential game analysis to both monopolistic and competing payment networks in an effort to examine the impacts of multihoming. They conclude that a greater tendency for end users to engage in multihoming makes it more likely that networks can co-exist but that the steady-state market shares of competing networks can diverge considerably depending participants propensities to engage in multihoming. They also argue that network distributors, such as issuers, play a crucial role in maintaining and expanding a network. Sun and Tse's analysis provides some measure of support for the Department of Justice's actions charging Visa and MasterCard with antitrust violations for their rules forbidding banks from becoming issuers of American Express and Discover cards.

Summary: Banking Antitrust

- Key rationales for bank mergers and acquisitions are potential cost efficiency gains and/or increased revenues owing to greater market power, either or both of which boost profitability of the consolidated institutions. Perceived diversification benefits that help reduce overall riskiness of the consolidated institutions also may spur merger and acquisition activity in banking, although there is mixed evidence regarding this motivation.

- Evaluating the social welfare effects of bank mergers involving institutions in perfectly competitive markets entails determining whether the trade-off between cost-efficiency gains and deadweight losses favors the former and hence generates welfare improvements. If markets are initially imperfectly competitive rather than perfectly competitive, market-power-enhancing effects of mergers result in banking consolidations being less likely to generate improvements in social welfare. Empirical research on impacts of mergers suggest that gains in both cost efficiency and market power do often arise simultaneously, implying that the social welfare effects of specific bank mergers or acquisitions must be evaluated on a case-by-case basis.
- U.S. bank merger policy is derived from general antitrust guidelines applied to nonfinancial firms but includes features specific to the market and regulatory environment that banks confront as financial firms. Maximum allowed bank concentration thresholds are adapted versions of critical Herfindahl–Hirschman index (HHI) values computed within regulator-defined relevant markets prior to and following a proposed merger. U.S. banking regulators can block proposed bank mergers or acquisitions, but even if banking regulators grant approval the U.S. Department of Justice can at its own discretion conduct a separate evaluation and challenge proposed consolidations.
- A fundamental issue relating to bank merger policy is the determination of the relevant banking market for application of concentration screens applied by bank regulators and antitrust officials. Traditionally, regulators have used metropolitan statistical areas as geographic proxies for relevant banking markets. Recent research, however, suggests a potential geographic broadening of U.S. banking markets. Some research also questions whether the Herfindahl–Hirschman index is the appropriate measure of potential market power in banking.
- Banks are heavily involved in transmitting payments, and in this role they participate in payment networks. Much recent work has aimed to apply industrial organization research on the economics of two-sided markets to these payment networks. This body of research suggests that in light of the complex balancing act entailed in balancing competing interests of card issuers, payment acquirers, cardholders, and retailers in the presence of network externalities, governmental regulation of card payment networks' fee structures is unlikely to yield welfare improvements and in fact could yield significant inefficiencies. It also generally concludes that welfare-improving competition among card payment networks is best promoted by permitting card issuers, payment acquirers, cardholders, and retailers to participate in multiple card payment networks. The implication is that recent antitrust challenges of restrictive rules imposed by some card payment networks probably have been appropriate.

Chapter 6
Bank Competition, Stability, and Regulation

Banks are among the most heavily regulated and supervised institutions on the planet. In addition to enforcing antitrust rules specially tailored to the banking industry, over the years, governments have implemented a wide range of regulatory policies, including restrictions on market entry, exit, and branching; explicit limits on bank loan and deposit rates; and a variety of rules governing the structure of banks' balance sheets.

Furthermore, designated banking agencies—in the United States, the Federal Reserve, the Office of the Comptroller of the Currency, the Federal Deposit Insurance Corporation, and Office of Thrift Supervision, and the National Credit Union Administration—subject banking institutions to periodic supervisory examinations to ensure adherence to such restrictions. U.S. banking supervisors conduct on-site examinations of banks to establish *CAMELS* ratings, where *CAMELS* is an acronym representing six components of supervisors' assessments of the banks' conditions: *C*apital adequacy, *A*sset quality, *M*anagement, *E*arnings, *L*iquidity, and *S*ensitivity to market risk. *CAMELS* ratings in each category are assigned on a scale from 1 to 5. A bank receiving ratings of 1 or 2 offers few sources of concern to supervisors, whereas a bank receiving ratings of 3, 4, or 5 presents moderate to significant degrees of supervisory concern. Curry et al. (2008) examine U.S. banking data for the 1985–1993 and 1994–2004 periods and find that *CAMELS*-rating downgrades generally do not affect a bank's overall loan growth but do tend to generate short-term reductions in higher-risk commercial and industrial loans.

This chapter reviews the rationales that have motivated construction of the massive apparatus for implementing regulation and supervision of banking institutions. In addition, it discusses bank efforts to rein in their risks via screening and monitoring activities, and it reviews theories and evidence regarding the effects of market structure on banking fragility. Finally, it describes key aspects of banking regulation that constrain the activities of individual banks and impinge on the structure of the banking industry.

D. Van Hoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-642-02821-2_6, © Springer-Verlag Berlin Heidelberg 2010

Banks as Issuers of Demandable Debt

A key advantage of firm-theoretic models of banking over portfolio-management models is the explicit consideration of decision-making with respect to both sides of a bank's balance sheet. Basic banking models surveyed in the previous chapters have not, however, explicitly considered risk and informational considerations that are inherent in banking.

One key source of banking risk is on the liability side of banks' balance sheets. Banks issue deposits, which function both as a form of indebtedness for funds they utilize to create assets and as embodiments of liquidity for depositors. Thus, banks issue demandable debts that are subject to depositor withdrawals at any time. Issuance of these debts exposes banks to the potential for "runs," such as those experienced by many U.S. banks at various points during the early 1930s (see, for example, Wicker, 1996).

The Diamond–Dybvig Model

One of the more influential analyses of the phenomenon of bank runs is the one-good, three-period (denoted 0, 1, and 2) model proposed by Diamond and Dybvig (1983). To develop an essential understanding of their basic framework, let's consider Dowd's (1992) exposition. All agents in the model, who are assumed to be risk-averse and are distributed over a unit interval, possess the same production technology permitting conversion of each unit of the good invested in period 0 into a unit of output if liquidated in period 1 and $R > 1$ units of output if held in place until period 2, where R is known with certainty. Alternatively, agents can simply store their goods rather than engage in investment, but if they do so the net return is always zero each period. Although agents are identical in period 0, at the beginning of period 1 they are revealed to be of two types: Type 1 agents, who are of this type with known probability p and who desire to liquidate their investments at the end of period 1, and type 2 agents who wish to hold their investments in place until the conclusion of period 2.

If $u(c_{i,j})$ is the utility derived from consumption in period i by agent of type j, then in period 0, prior to realization of the risk of being either type 1 or type 2, agents can engage in a risk-sharing arrangement via a form of insurance contract involving $c_{1,1} > 1$ (type 1 agents being able to consume more than a unit of the good in period 1) and $c_{2,2} < 1$ (type 2 agents consuming less than the full gross investment return when they engage in consumption in period 2).

An Optimal Risk-Sharing Contract

Diamond and Dybvig demonstrate the existence of an optimal risk-sharing contract optimal risk-sharing contract satisfying three conditions: (i) $c^*_{1,2} = c^*_{2,1} = 0$ (that is, type 1 agents cannot possibly derive utility in the second period and hence do not consume then, and type 2 agents do not consume early), (ii) $u'\left(c^*_{1,1}\right) = \rho R u'(c^*_{2,2})$

where ρ is the subjective rate of time preference with $\frac{1}{R} < \rho \leq 1$ (thereby assuring appropriate alignment of marginal utility with marginal productivity), and (iii) $pc^*_{1,1} + (1-p)\dfrac{c^*_{2,2}}{R} = 1$ (satisfaction of the constraint that one unit of the good is available for use ex ante by an agent). Because the second condition implies that $\rho R > 1$, it follows that $\dfrac{u'(c^*_{1,1})}{u'(c^*_{2,2})}$ exceeds unity, and the marginal utility derived from a type 2 agent from consuming in period 2 is less than the marginal utility that a type 1 agent experiences from consuming in period 1, requiring $c^*_{2,2} > c^*_{1,1}$.

Thus, under this optimal risk-sharing contract, after a type 2 agent learns of her random assignment to that type, she has no incentive to try to pass herself off as a type 1 agent, because $c^*_{2,2} > c^*_{1,1}$. At the same time, a type 1 agent also has no incentive to try to pass himself off as a type 2 agent in light of the fact that $c^*_{1,1} > 1 > c^*_{1,2} = 0$. Hence, in principle the contract is incentive-compatible.

The Diamond–Dybvig Intermediation Solution and the Problem of Runs

In spite of the theoretical existence of an incentive-compatible contract, mutual insurance arrangement in which agents simply announce their types cannot be implemented, because each agent could claim to be of type 1 just to get the insurance handout in period 1. Diamond and Dybvig suggest that an intermediary could implement the optimal contract by taking in deposits, investing them in the production process, and promising depositors a payment of $r_1 = c^*_{1,1} > 1$ if they engage in withdrawals in period 1, provided that the intermediary has assets available to liquidate.

Diamond and Dybvig assume that withdrawal demands arrive randomly during period 1 and that the intermediary honors them sequentially—that is, intermediaries satisfy an exogenously specified sequential service constraint. Under this presumed structure, Diamond and Dybvig show that the intermediary's deposit contract can support the optimal risk-sharing contract as a Nash equilibrium. The equilibrium is not unique, however. An alternative equilibrium is a "run" equilibrium in which all of the intermediary's assets are liquidated in period 1 as a consequence of a "panic" by type 2 depositors worried—due to some unmodeled source of external uncertainty—that the bank will run out of assets prior to period 2 when they are scheduled to consume goods. Under the assumption that the intermediary knows the value of p, the share of agents that will be type 1, the intermediary can modify its deposit contract to prevent the run equilibrium. The intermediary can accomplish this via a "suspension-of-convertibility contract": It can promise to redeem deposits on demand in period 1 until that fraction of agents has received funds and then suspend further convertibility of assets until period 2, thereby assuring optimal returns in period 2 and ruling out a reason for type 2 agents to possess sufficient "worry" to launch a run.

Diamond and Dybvig demonstrate, however, that a suspension-of-convertibility contract cannot prevent a run if p is a random variable whose realization is determined at the beginning of period 1 and hence is unknown in period 0. Diamond and Dybvig argue that in this circumstance, an outside agency, presumed to be costlessly operated by the government and perhaps called a "deposit insurer," could intervene. It could do so by guaranteeing—via an appropriately designed taxation-and-subsidization scheme—that after the fact, those who withdraw in period 1 will receive their optimal consumption bundle contingent on the ex post realization of p. This policy removes the incentive for type 2 depositors to engage in a run, because they can rest assured that the agency's guarantee will yield the optimal risk-sharing contract's consumption package across both periods, including period 2 when type 2 agents are assigned to consume.

Evaluating the Diamond–Dybvig Analysis

The Diamond–Dybvig model possesses a number of attractive features. First, it motivates how a financial intermediary might come into existence—namely, as a mechanism for implementing a risk-sharing contract among agents with differential timing of consumption. Second, it provides a rationale for bank runs as observed breakdowns in the optimal risk-sharing arrangement owing to worries by depositors that Diamond and Dybvig suggest might be generated by knowledge of weak intermediary earnings on assets owing to economic shocks, observation of a run on another intermediary, or some other outside event. Third, the model offers a possible motivation for real-world governments to become involved in operating deposit-insurance schemes.

Nevertheless, work by Jacklin (1987) and Wallace (1989) indicates that the Diamond–Dybvig framework requires imposing additional background assumptions in order to justify the existence of a financial intermediary. Jacklin shows that Diamond and Dybvig's intermediation mechanism for ensuring the optimal risk-sharing contract could be replaced by an alternative mechanism in which agents invest in a mutual fund in period 0, which in turn issues agents equity claims that mature in period 2 and upon which the mutual fund pays dividends. The mutual fund structures these dividends in a way that ensures that only type 1 agents will desire to sell off their equity shares in period 1 and only type 2 agents will wish to buy those shares. Effectively, agents engage in trade between periods 1 and 2, and determination of a market clearing price of shares effectively permits implementation of the optimal risk-sharing contract without a necessity for appealing to a Diamond–Dybvig-style intermediary. Wallace offers a means of rescuing the Diamond–Dybvig-style intermediary by suggesting that their framework applies to an arguably believable setting in which agents are "isolated" in the sense in which trade across periods 1 and 2 is sufficiently costly that agents will prefer to deal with an intermediary instead. Then when p is unknown ex ante, the Diamond–Dybvig "run" equilibrium can emerge once more and potentially can be ruled out via government intervention.

As Dowd (1992) notes, however, the Diamond–Dybvig intermediary is not quite recognizable as a "bank." Real-world banks promise fixed rates of return to depositors in all periods, but Diamond–Dybvig type 2 depositors are, like owners of equity shares, residual claimants. Thus, their "bank deposit contract" is actually something of a hybrid debt-equity instrument rather than a demandable deposit as commonly utilized in banking markets.

Another difficulty is that intervention in the form of "deposit insurance" by a government agency can create moral hazard problems that must be addressed by additional, likely non-costless, interventions. After all, the expected value of an agency bailout subsidy to any given intermediary would increase with the probability of default. This probability would increase with the promised return to withdrawals in period 1, which in turn would rise as competitive intermediaries would seek to maximize the value of the subsidy by aggressively offering higher deposit rates. As discussed by Dowd, this recognition has led to suggestions that the Diamond–Dybvig "deposit insurance" scheme either would have to be supplemented with additional regulation, such as deposit rate ceilings (Anderlini, 1986) or reserve requirements (Freeman, 1988) or replaced by a private insurance mechanism (Dowd, 2000).

Calomiris and Kahn (1991) add to the list of concerns about implications of the Diamond–Dybvig framework. They point out that in Diamond–Dybvig-style models, suspension of convertibility typically emerges as preferable to bank closure, yet historically the latter has been more likely to occur. In addition, they note that the Diamond–Dybvig framework takes as given the sequential service constraint for bank withdrawals, even though this first-come, first-serve rule for payments by banks to depositor-creditors contrasts with standard restraints on troubled firms facing bankruptcy. Calomiris and Kahn propose the existence of bank deposits as demandable debts that provide an incentive-compatible solution to the problem of balancing the comparative advantage of banks in allocating credit against the ability of banks to act contrary to the interests of uninformed depositors.

Until recently, there has been relatively little empirical evidence brought to bear to test implications of the Diamond–Dybvig analysis of bank runs, primarily because insufficient data have been available. A recent contribution by Iyer and Puri (2008) is an interesting exception, however. These authors examine data from one of several cooperative banks that experienced runs in March 2001 in the Indian state of Gujarat. They find evidence, consistent with the prediction of the Diamond–Dybvig analysis, that the presence of deposit insurance helps limit the potential for bank runs. Nevertheless, Iyer and Puri also find that another potentially important factor is social network effects that influence the speed of contagion of bank runs. They identify two other crucial factors, which are the length and depth of the bank-depositor relationship. Depositors who have had a long-term customer relationship are less likely to participate in a run on the bank. In addition, depositors who are also borrowers from the bank are less likely join in a bank run.

Iyer and Puri's conclusions point to important features absent from models in the Diamond–Dybvig mold. One is an absence of contemporaneous depositor heterogeneities. Diamond–Dybvig-style models typically allow for hetergeneities across

time, such as the Diamond–Dybvig assumption that depositors are endowed with predispositions to withdraw funds in different periods. Otherwise, however, depositors are identical in typical Diamond–Dybvig-style models. Another usually missing feature is a meaningful depiction of the asset side of a bank's balance sheet. In real-world situations in which bank runs occur, depositor "worries" are often based in concerns about risky asset return realizations influenced by conscious decisions of bank managers.

Banks as Screeners and Monitors

Over more than two decades, based largely on work by Diamond (1984, 1991, 1996), the banking literature has focused considerable attention on a characteristic of banking that allegedly makes banks "special" (see Kwan, 2001). This characteristic is banks' role in screening and monitoring loans to contain the highly idiosyncratic risks associated with these financial assets. Banks performing these task more efficiently using existing resources presumably make higher-quality loans and thereby reduce overall risks of loss and failure probabilities. Recently, Diamond and Rajan (2000, 2001, 2006) have built on these ideas to propose theories of the functioning of banking markets and of the role of banks as issuers of liabilities that function as money.

Diamond (1984) highlights how exposure to the moral hazard behavior of borrowers and thus the potential for loan defaults makes banks' activities as delegated monitors a key function. Separately, Diamond (1991) analyzes the case of endogenous bank monitoring. He does so under the assumption that all lenders face fixed and identical cost of monitoring, hence allowing for no cross-sectional distribution of monitoring and non-monitoring lenders. Consequently, all banks in Diamond's framework either do or do not engage in loan monitoring.

In real-world banking systems, however, it is arguable that banks are not necessarily equally attentive to risks in their loan portfolios. Indeed, Carletti et al. (2007) suggest that in cases in which banks lend to borrowers engage in relationships with multiple banks, some banks may free-ride on monitoring efforts of others. This effect that potentially can predominate over positive monitoring incentives generated by diversification benefits derived from the wider range of loans that result, resulting in less monitoring by some banks. In contrast, for other banks the consequence can be higher levels of monitoring than otherwise would occur.

Evidence on Bank Monitoring Activities

Most evidence regarding whether banks provide special monitoring services is indirectly derived from loan announcement effects, bond yields and rates on syndicated loans, and loan sales. Direct evidence relating to bank loan monitoring has recently emerged only in a single study, Mester et al. (2005), which is discussed later in this section.

Evidence from Bank Loan Announcement Effects

A comprehensive examination of evidence about whether banks perform special screening and monitoring functions has been provided James (1987), who analyzes more than 200 firms' financing announcements—including announcements of both private and public debt placements and loans from banks—between 1974 and 1983. James finds that announcements of debt issues tended to generate abnormal negative stock returns for the borrower. In contrast, announcements of bank loans generate positive abnormal returns. This, James concludes, is evidence that valuable private information is encompassed with the lender-borrower relationship, which suggests that banks do indeed provide screening and monitoring functions with perceived value to the recipients.

Building on James (1987), Billett et al. (1995) construct and examine a sample of nearly 1,500 corporate loan announcements between 1980 and 1989 and examine whether the identity of the lender affected the market response of the borrower's stock returns. They find that firms borrowing from banks with higher credit ratings experienced higher abnormal returns, with each one-unit increase in the lender's credit rating boosting the borrower's return by 20 basis points during the day following a loan announcement. Billett et al. propose that a higher positive response of a borrower's stock returns to announcement of a loan by a lender with a superior credit rating likely reflects loan-announcement signaling effects. They note, for instance, that a stronger credit rating for a lender could proxy for that bank's monitoring effectiveness, which in turn influences the perceived likelihood that it will repay depositors and other creditors. To the extent that this signaling effect lies behind the higher positive response of a borrower's stock returns, Billet et al.'s results suggest that markets perceive that banks have heterogeneous monitoring capabilities that, in turn, provide an indication of the quality of their customers' loans.

Coleman et al. (2006) develop a proxy measure of the level of labor input into loan monitoring based on the share of salary expenses to total non-interest expenses, which they show using data on more than 1,000 U.S. banks in latter 1990s is directly related to both loan maturity and to the loan yield spread. Lee and Sharpe (2009) investigate the relationship between this monitoring measure and loan announcements and find a relatively small but significant effect on bank returns.

Evidence from Firm Investment and Bond Yields

As an alternative way of gauging how the use of bank lending services affects borrowers' performances, Hoshi et al. (1991) examine panel data on Japanese manufacturing firms. Hoshi et al. find that firms that are members of industrial groups, or *keiretsu*, receiving financing from large city banks are able to undertake more investment than non-members because of relaxed liquidity constraints associated with closer *keiretsu* ties with banks. Although they acknowledge that various services provided by keiretsu contribute to this outcome, they suggest that a key factor accounting for this outcome could be the loan-monitoring activity of the *keiretsu* banks.

Datta et al. (1999) study yield spreads on public bond offerings of firms that are bank borrowers. They find that bonds issued by firms with longer-standing borrowing relationships with banks have yield spreads about 68 basis points lower than bonds issued by firms without such relationships. This finding, the authors suggest, supports Diamond's (1991) thesis that market participants view bank monitoring functions as value-enhancing for borrowers.

Evidence from Syndicated Loans and Loan Sales

Casolaro et al. (2003) examine more than 14,000 syndicated loan facilities organized by a single lending institution between 1990 and 2001. They find that for syndicated loan facilities created for all but the largest, most reputable borrowers that presumably have most market transparency, there is a negative relationship between the loan rate granted by lending syndicates and the share of credit risk retained by the original facility organizer. Although this relationship might also be explained by a rising marginal cost associated with greater syndicate participation, Casolaro et al. (2003) interpret this result to be consistent with a certification effect. They suggest that greater retention of credit risk by a syndicated loan organizer increases that lender's incentive to maintain its level of monitoring, thereby reducing the interest rate at which the syndicate as a whole is willing to extend credit to the borrower.

Gande et al. (1997) likewise find evidence of a certification effect that has operated when banks' section 20 subsidiaries underwrite firm debt securities. Yields on such securities issued between 1993 and 1995, they find, tended to be lower than those underwritten by traditional investment banks, indicating that markets assigned value to the monitoring functions performed by banks in their dealings with firms in lender-borrower relationships.

Kim et al. (2005) examine loan loss provisions from a panel of more than 100 Norwegian banks over the 1993–1998 interval. They find evidence consistent with the interpretation that borrowers use credit relationships with higher-quality, lower-loss banks to signal their creditworthiness to other parties.

Gorton and Pennacchi (1995) study the characteristics of a sample of major U.S. banks that engaged in loan sales in 1987 and 1988 and examine the features of the 872 loans these banks sold. They observe that the banks typically retained a smaller share of the higher-risk loans that presumably necessitate greater monitoring, which their theoretical analysis indicates reduces the loan default premium that loan buyers require in order to purchase the loans. To induce borrowers to purchase higher-risk loans at prices incorporating lower default premia, banks retained larger portions of these higher-risk and presumably more "special" loans for which greater monitoring effort was required.

Dahiya et al. (2003) conduct an event study focused on the relationship between secondary-market bank loan sales that were announced between 1995 and 1998 and the subsequent performances of borrowers whose loans were sold. They find evidence that loan sales are associated with subsequent defaults and consequent lower returns on the loans, which also supports the existence of a certification effect. In addition, Dahiya et al. provide evidence both that the sale of loans foreshadowed

ultimate failure by a significant portion of the affected borrowers and that such failures could not have been readily predicted using public information. These conclusions, Dahiya et al. suggest, offer support for the hypothesis that banks are able to detect loan risks not readily observable to other investors.

Dahiya et al. also investigate the characteristics of banks that engaged in loan sales during their sample period. They find that banks that sold loans during their sample period were distributed relatively evenly among other non-selling banks based on income-asset ratios and that selling banks' returns were unaffected by the loan sales. Nevertheless, most banks that sold loans were relatively low-capital banks that rated among the 50 percent of banks with bad loan reserves in relation to total assets. By and large, the loan sales generated increases in these banks' capital ratios, suggesting that for these banks, the sales entailed efforts to shrink loan portfolios and reduce asset risk.

Direct Evidence of Bank Monitoring Activities

The studies discussed above provide only indirect evidence regarding the hypothesis that banks uniquely provide services as monitors that improve the quality and, hence, the market value of loans. The best sources of data regarding resources that banks apply to the task of monitoring and the nature of bank monitoring operations are banks themselves. Such data are typically proprietary, however. Economists' lack of access to such data has tended to limit the capability to directly evaluate the scope of banks' monitoring activities.

One exception is a recent study by Mester et al. (2005), which examines data from a single Canadian bank. Mester et al. find that monitoring of transactions deposits held with the lender by a borrower—conducted via item-by-item reconciliations of transactions—enables a bank to verify the reliability of the borrower's statements about flows of accounts receivables and inventories.

The analysis of Mester et al. indicates that the provision of transactions deposits to borrowers permits banks to access information about borrowers' behavior that is not available to other potential lenders. Mester et al. suggest that finance companies or other lenders, including banks, that lack direct depository relationships with borrowers must expend additional resources to attempt to partially compensate for absence of such information. Furthermore, even banks with access to depository transactions data must expend resources to utilize it for purposes of monitoring the activities of borrowers.

A Monitoring Model with Heterogeneous Banks

To contemplate the implications of accounting for bank monitoring choices, consider the following version of the model developed by Kopecky and VanHoose (2006). In this framework, monitored loans have a lower loan default rate than non-monitored loans. Aggregate lending by the banking system as a whole depends on

the monitoring decisions of all the individual banks. As a consequence, the market loan rate and the equilibrium share of banks that monitor are simultaneously determined.

Behavior of Monitoring and Nonmonitoring Banks

Suppose that, as in Elyasiani et al. (1995) banks are perfectly competitive, but suppose that the banking environment is static. The profits of bank i, which under assumption of an absence of nondeposit liabilities faces the balance sheet constraint, $L^i + S^i = (1 - q)D^i$ are given by

$$\hat{\pi}^i = \hat{R}_L^i L^i + r_s S^i - r_D D^i - \frac{\alpha}{2}(L^i)^2 - \frac{\beta}{2}(S^i)^2 - \frac{\theta}{2}(D^i)^2 - \left(\frac{i}{1-i}\right)\frac{c}{2}(L^i)^2,$$

where the parameters α and θ are nonnegative constants governing the magnitude of quadratic resource costs of lending and servicing deposits, and where $\left(\frac{i}{1-i}\right)\frac{c}{2}(L^i)^2$ is the bank's marginal loan-monitoring cost—*if* the bank chooses to monitor its loans, as discussed below—where c is a positive constant, with $c \leq \alpha$, so that monitoring costs at the median bank ($i = 1/2$), are no higher than other resource costs associated with lending.

Banks are distributed along a unit interval according to their monitoring costs, which are assumed to be monotonically increasing. Thus, bank $i = 0$ incurs the lowest (zero) marginal loan-monitoring cost, and bank $i = 1$ incurs the highest marginal loan-monitoring cost. This variation in marginal loan monitoring costs across banks is intended to capture the potential for management skills to differ across banks. At the same time, the monitoring-cost component c reflects common technological aspects of the loan-monitoring processes utilized by banks. Note that the form of this loan-monitoring cost function implies that there is an increasing *internal* marginal costs of monitoring loans at bank i: The greater is the volume bank loans at the bank, the higher is the marginal monitoring cost that it incurs.

In addition, the function $\left(\frac{i}{1-i}\right)\frac{c}{2}(L^i)^2$ implies rising *external* marginal monitoring costs. The derivative of the assumed monitoring cost function with respect to i is $\left(\frac{1}{(1-i)^2}\right)\frac{c}{2}(L^i)^2 > 0$. Consequently, the ith marginal bank that engages in monitoring has a higher marginal cost of monitoring loans than the more efficient banks located below i on the unit interval, implying that the banking system as a whole faces rising external monitoring costs. In contrast to Diamond (1991), therefore, monitoring is not a fixed-cost activity at the level of either the individual bank or the industry as a whole.

A bank may choose to monitor its loans because borrowers are assumed to use a portion of loan proceeds to finance unproductive activities that results in a loan default that yields no remaining liquidation value. Following Kopecky and VanHoose (2006), let's suppose that a monitoring bank is *always* successful in preventing such unproductive uses of funds, in which case the proportion of non-

defaulting loans is η (so that the loan default rate is $1-\eta$), which is based on macro factors that banks and borrowers cannot influence. Hence, the effective return to lending for a bank that engages in monitoring and incurs the marginal monitoring cost $\left(\frac{i}{1-i}\right)\frac{c}{2}(L^i)^2$ is equal to $\hat{R}_L^M = \eta r_L$. If a bank opts not to monitor loans, it incurs no monitoring cost but also exposes itself to the potential for a reduction in its loan payoff rate equal to δ where $0 < \delta < \eta$, implying a loan default rate equal to $1+\delta-\eta$ and effective return to lending at a bank that does not monitor loans equal to $\hat{R}_L^{NM} = (\eta-\delta)r_L$.

Loan Market Equilibrium and Equilibrium Monitoring

The profit-maximizing lending by a monitoring bank is

$$L_M^i = \frac{[\beta(1-q)^2+\theta]\eta r_L - \theta r_s - \beta(1-q)r_D}{[\alpha\beta(1-q)^2+\theta(\alpha+\beta)] + c\left(\frac{i}{1-i}\right)[\beta(1-q)^2+\theta]}, 0 \le i \le \Omega.$$

As i increases and the marginal monitoring cost of bank i rises, the ith bank's level of lending declines. Let bank Ω be the bank on the external margin with regard to loan monitoring, so that non-monitoring banks are in the interval $\Omega < i < 1$. A non-monitoring bank maximizes profits with $c = 0$ but with $\hat{R}_L^{NM} = (\eta-\delta)r_L$, which yields identical loans at each of the non-monitoring banks:

$$L_{NM}^i \equiv L_{NM} = \frac{[\beta(1-q)^2+\theta](\eta-\delta)r_L - \theta r_s - \beta(1-q)r_D}{\alpha\beta(1-q)^2+\theta(\alpha+\beta)}, \Omega < i \le 1,$$

For the banking industry, therefore, market loan supply (L^S) is the total of all loans by monitoring and non-monitoring banks, given by

$$L^s = \int_0^{\Omega} L_i^M di + (1-\Omega)L_{NM}$$

Kopecky and VanHoose consider a linear functional form for the public's loan demand, given by $L^d = l_0 - l_1 r_L$. Equalizing the quantity of loans supplied and the quantity of loans demanded determines values of the market loan rate for any given share of monitoring banks, Ω, given by the loan market equilibrium (LME) locus in Fig. 6.1.

The bank on the external margin is indifferent between monitoring or not monitoring when profits are equalized. Substituting the optimal quantities of loans as well as securities and deposits (which can be solved from the model) into the relevant profit functions for monitoring and non-monitoring banks and equating the results yields a second relationship between the loan rate the share of monitoring banks, given by the equalized-profit (EP) locus in panel (a) of Fig. 6.1.

In fact, the general forms of the LME and EP schedules are determined by complex polynomial relationships implied by the expressions above. Kopecky and VanHoose utilize calibrated simulations based on U.S. banking data that indicate that the LME schedule slopes upward—that is, as the share of banks that incurs the

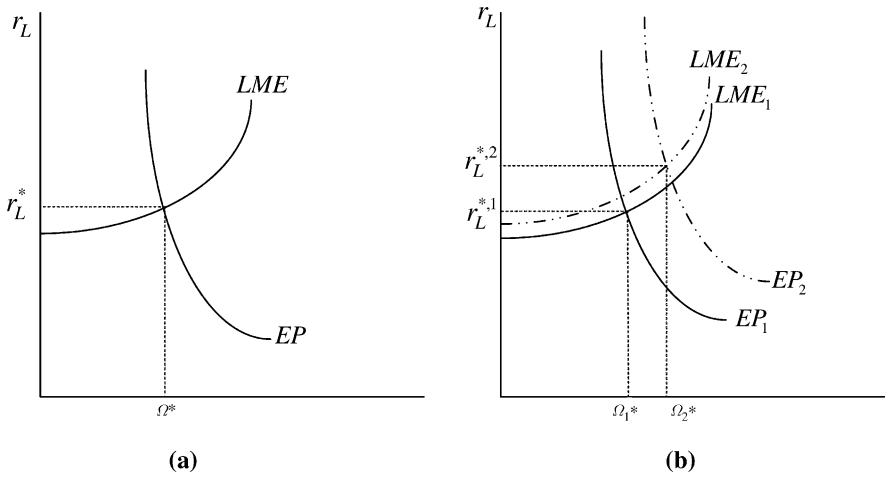

Fig. 6.1 Joint determination of the market loan rate and the equilibrium share of monitoring banks

cost of monitoring loans increases, market loan supply decreases, hence the equilibrium loan rate rises. In principle, the EP schedule could slope either upward—as the loan rate rises and revenues increase, *ceteris paribus*, more banks might be willing to incur the cost of monitoring their loans—or downward—as the loan rate rises, the smaller size of the marginal bank's loan portfolio may generate revenues that fail to cover monitoring costs. The calibrated simulations indicate that the latter effect predominates, hence the depicted downward slope of the EP locus.

Panel (b) of Fig. 6.1 displays the effect of changes in key parameter values. For instance, an increase in the value of δ and hence the loan default rate reduces bank loan supply and hence boosts the market loan rate for any given share of monitoring banks, resulting in an upward shift in the LME schedule. At the same time, a rise δ raises profits for monitoring banks relative to those that do not monitor, which causes the share of monitoring banks to increase at any given loan rate, thereby shifting the EP schedule rightward. Based again on calibrated simulations, the effects are increases in both the market loan rate and the equilibrium share of banks that monitor their loans. Analogous effects to those depicted in panel (b) follow if there is a reduction in overall monitoring costs caused by a rise in the value of c. Thus, more banks monitor their loans if the banking industry is confronted with greater risk of loan default or experiences an improvement in loan-monitoring efficiency.

This analysis suggests that bank loan monitoring activities typically must impinge on market outcomes in the banking industry. Conversely, realized market outcomes also influence the extent to which banks monitor their loans. It follows that regulations aimed at enhancing the safety and soundness of the banking industry will exert dual effects, as discussed in detail in Chapter 7.

The Relationship between Banking Competition and Risks

Is bank entry regulation that makes banking markets more concentrated—and potentially less competitive—than they otherwise might be effective as means of lowering the riskiness of bank asset portfolios, decreasing insolvency risk, and reducing the potential for bank runs? Ultimately, answering this question requires uncovering the relationship between banking stability and competition.

Perfect Competition and Bank Risks

As discussed in Chapter 3, the efficiency properties of perfectly competitive banking markets are socially desirable. Nevertheless, several arguments have been advanced suggesting that an undesirable property of perfect competition in banking is that this market structure also promotes excessive risk taking by individual banks and, consequently, instability of the banking system as a whole.

The "Excessive Deposit Competition" Argument

There is a longstanding argument suggesting that unhindered competition in deposit markets sows the seeds of banking instability. According to this argument, which has variously been advanced throughout history, excessive competition for deposits induces banks to pay deposit rates that are "too high," because a portion of funds attracted at market interest rates will be prone to greater risk of unexpected withdrawal that can fuel a bank run.

In an application of the Diamond–Dybvig (1983) model, Smith (1984) formalizes this argument by considering a setting in which depositors know the probabilities of their withdrawals but banks do not. When banks offer deposit contracts providing profitable rates of return, an equilibrium non-existence result emerges, implying a lack of viability of the banking system. Other work in this vein includes Goldstein and Pauzner (2005) and Rochet and Vives (2004). These analyses are also modified versions of the Diamond–Dybvig framework, in which fundamental determinants of rates of return are stochastic and agents receive noisy signals of these fundamentals. As a consequence, sudden changes in values of fundamentals or in expectations of fundamentals can induce bank runs. Although such runs can be "efficient" if the bank's long-run market value is below its current liquidation value, inefficient runs in which the reverse is true can also occur. Furthermore, the probability of such inefficient runs is higher if banks offer higher returns to depositors.

Matutes and Vives (1996, 2000) develop models that explicitly couple the potential for banking collapse with imperfect deposit-market competition via offerings of differentiated products by banks. In Matutes and Vives (1996), depositors' ex ante failure perceptions influence the degree of rivalry and hence ex post equilibrium outcomes and hence the actual probability of failure. A key consequence is that the implementation of a deposit insurance scheme has ambiguous consequences. In Matutes and Vives (2000), socially excessive deposit rates and deposit holdings

emerge as equilibrium outcomes with or without deposit insurance, with welfare implications contingent on the intensity of competition.

One commonly proposed solution to the alleged tendency for excessive competition to contribute to triggering runs is deposit rate ceilings such as those put into place in the United States by the 1933 Glass-Steagall Act. Indeed, this is a suggestion offered by Smith (1984) and Matutes and Vives (2000). A problem with basing support for interest-rate controls on the excessive-competition argument, however, is that theoretical analyses providing the logic for these controls typically fail to account for the existence of alternative saving vehicles to deposits. In virtually all applications of the Diamond–Dybvig analysis, depositors essentially are captive savers. If alternative saving vehicles are available as imperfect but relatively close substitutes to bank deposits, however, deposit-rate regulation can itself create a foundation for banking instability: Failure to index deposit rate ceilings to market interest rates can expose to a wave of disintermediation—essentially a slow-motion and potentially highly inefficient bank run—if there is a sudden rise in market rates above controlled deposit rates. Hence, the imposition of deposit-market restrictions arguably could create conditions giving rise to a chronic stream of welfare losses with a discounted present value potentially exceeding that of the acute losses caused by bank runs.

The Competition-Illiquidity Argument

Even if not "excessively competitive," does the nature of perfectly competitive banking markets expose banks to greater risk of contagion? Allen and Gale (2004) offer a theoretical argument suggesting that perfect competition in interbank markets can contribute to generalized instability. This argument hinges on the atomistic-agent property of a perfectly competitive market. If each bank individually has no impact on the market equilibrium, then there is no incentive for other banks to lend to any given bank experiencing liquidity problems. Nevertheless, if liquidity distress at some banks gives rise to spillover effects that create liquidity problems for other banks, the banking system as a whole would be better off by lending to illiquid banks. Yet there is no mechanism to induce such lending. Hence, there is an unresolved coordination failure resulting from atomistic competition coupled with market failures arising from third-party spillovers across institutions.

Essentially, this particular argument advanced by Allen and Gale formalizes Goodhart's (1988) rationalization for the creation of central banks. Goodhart argues that a central bank provides a mechanism for coordinating liquidity management that benefits all institutions. Because it is not in an individual bank's interest, ex ante, to favor creation of such an institution when its operations are funded by explicit or implicit taxation of the private banking system—even though is existence could prove useful ex post—there is a potential role for government to establish such an institution as an independent arbiter and head of a "club" of banks.

Nevertheless, as discussed by Dowd (1994), a number of theoretical counterarguments can be advanced, leading to the opposite conclusion that a freer market is more likely to provide a stable banking environment. More recently, De Vries (2005)

analyzes a theoretical model of systemic risk arising from linkages in the interbank deposit market and concludes that concentration of risk within individual institutions leads to more frequent isolated failures. Segregation of risk across multiple institutions reduces the likelihood of systemic risks, suggesting that a more competitive interbank market actually could be more stable than a more concentrated market.

The Competition-Asset Risk Argument

Ogura (2006) offers an alternative foundation for why competitive banking might generate instability by influencing individual banks' and the aggregate banking system's asset risk via a "monkey-see-monkey do" mechanism. In his theoretical framework, a bank uses newly arriving information to engage in Bayesian updating of its belief about the probability of investment success by a prospective borrower, including information about whether the loan applicant has successfully received credit from a rival institution. If a rival extends a loan to the prospective borrower, the bank raises its estimate of the probability of project success for that borrower, thereby making the bank more likely to reduce its credit standards and extend a loan to that borrower. Ogura's analysis suggests that the magnitude of the reduction in lending standards is reduced by a greater prevalence of relationship banking or a larger number of rival banks, which in his model reduces the marginal impact of a rival's loan on the bank's update of the probability of borrower success. Nevertheless, an increase in the number of banks leads to more borrowers obtaining loans. Furthermore, in good economic times as more borrowers qualify for loans there is a tendency for each bank to extend more loans as its rivals expand their lending. The result essentially is a herding effect, with banks collectively loosening their credit standards during economic booms—a result arguably consistent with stylized facts associated with banking during the 2000s. Ogura concludes, therefore, that a highly competitive banking system may require more regulatory supervision.

Acharya and Yorulmazer (2008) show that herding can also result when an individual bank seeks to minimize the effects resulting from news relating to its own portfolio risks that can adversely affect its cost of borrowing. In their model, the potential for such effects induces each bank to undertake investments that are correlated with those of its rivals, which helps prevent adverse news from pushing up borrowing costs relative to those of other banks. The result, naturally, can be greater exposure of the entire banking system to risks that ultimately will affect all banks similarly in light of their asset herding behavior.

Furthermore, Keeley (1990) argues that a reduction in a bank's charter value owing to increased competition reduces the incentive for a bank to take on greater asset portfolio risk. Consistent with this proposition, Marinč (2008) has provided an analysis suggesting that greater competition can reduce the incentive for each individual bank in a market to monitor its loans, an effect that Marinč shows can be at least somewhat mitigated by greater portfolio diversification.

In addition, Allen and Gale (2004) note that a zero-economic-profit equilibrium dictated by long-run forces of competition gives economic-profit-seeking banks

marginal incentives to take on greater asset risk as the intensity of competition increases. From this perspective, increased competition does not "cause" failures or runs, but greater competition expands the risk exposure of the entire banking system to adverse changes in fundamental values of assets.

It is conceivable, therefore, that there can be "too much" competition in the banking industry. The analyses of Ogura, Acharya and Yorulmazer, Keeley, and Marinč suggest that direct supervision of banks' asset risks may be appropriate. Allen and Gale's study indicates that regulations constraining entry and allowing banks to earn positive economic profits on less-risky portfolios of loans and other assets also might enhance stability of the banking system.

Market Power and Bank Risks: Theory and Evidence

The argument that restraints on competition among banks contributes to greater industry stability has long been offered as a rationale for bank entry regulations in the form of licensing or chartering requirements that limit entry into banking markets. Buttressing this argument is the fact that banks earning economic profits thereby tend to be cushioned from adverse changes in fundamental values of assets at a given degree of exposure to asset risk.

Such arguments are commonly encountered in the literature. Consider, for instance, the analysis of Hughes et al. (1999) discussed in Chapter 5, suggesting that risk diversification benefits of banking consolidation lead to improve financial performance on the part of merged banks. Hughes et al. argue that such consolidations are also socially beneficial in part because of a presumed enhancement of bank safety resulting from the greater market concentration which results. This presumption receives support, for instance, from recent work by Ariss (2008), who studies data from more than 800 banks in 60 developing nations between 1999 and 2005 and concludes that increased market power boosts cost and profit efficiency and hence bank stability. In addition, an examination of data from 69 countries from 1980 to 1997 by Beck et al. (2006) yields the conclusion that banking crises are less likely in nations with more concentrated banking systems.

Competition and Risk: Theory

As Allen and Gale (2004) discuss, from a theoretical standpoint there is no ironclad argument that banks with market power will necessarily choose assets with lower risks. As they point out, even when prospective borrowers and depositors regard banks' loans and deposits as homogeneous, the portfolio choices of banks depend on the presumed nature of oligopolistic rivalry. For instance, Martiez-Miera and Repullo (2008) examine a dynamic theoretical model that posits Cournot loan-market rivalry. Their model allows for banks that do not fail in a given period to finance new entrepreneurial projects in the next period, thereby generating an endogenous bank franchise value and giving banks an incentive to lend prudently.

Analysis of this model yields a predicted U-shaped relationship between competition and the risk of bank failure, implying that greatest probabilities of failure arise in either highly competitive markets or markets in which banks have greatest market power. In the context of their Cournot framework, therefore, the greatest risk of banking instability lies with intermediate degrees of bank market rivalry.

Allen and Gale also point out that differentiation of loans and deposits via non-price competition further complicates banks' strategic choices. A natural consequence is that this additional level of strategic decision-making impinges on bank risk choices, further muddying the competition-stability nexus.

A bank on the margin regarding its level of monitoring for moral hazard risks is also likely to face differential incentives depending on how much market power it possesses. Caminal and Matutes (2002) construct a theoretical framework in which banks finance investment projects of borrowers, which exposes banks to moral hazard problems that can be addressed either via monitoring or credit rationing. Firms undertake fewer projects in the face of a higher loan rate and more projects with increases in monitoring. Greater market power leads banks to boost loan rates and monitoring activities. Overall, Caminal and Matutes conclude, investment and hence potential project failures respond ambiguously. In a different theoretical framework aimed at assessing the simultaneous effects of bank competition on credit-market outcomes and equilibrium borrower bankruptcy risk, Koskela and Stenbacka (2000) likewise suggest that greater competition need not lead to greater credit risks.

There is yet another complicating factor to consider. In a review of the literature on the concentration-fragility nexus, Boyd and De Nicolo (2005) conclude that most work assumes that asset prices and return distributions are exogenous to the structure of banking markets. Of course, the risk characteristics of loans extended by a monopoly bank could well differ from those of loans extended by atomistic competitors. Boyd and De Nicolo suggest that taking into account endogeneities in asset prices and the distributions of returns strengthens the case for greater competition enhancing financial stability.

Bank Size, Competition, and Risk: Evidence

Is a more concentrated banking industry comprised of relatively large institutions more safe and sound? Berger et al. (2008), who examine data 1999–2005 data from more than 8,000 banks in 23 industrial nations, find evidence of a lower overall degree of risk exposure at banks possessing greater market power. Nevertheless, they simultaneously find that banks with more market power are exposed to greater credit risk. In a study of an unbalanced panel of 95 banks in the European Union, Stolz (2007) finds little evidence of a relationship between banks' charter values and their propensity to take on risks.

Boyd and Gertler (1993) document that during the 1991 recession, large U.S. banks accounted for a disproportionate amount of U.S. loan losses. Boyd and Graham (1996) show that from the early 1970s through the mid-1980s, large U.S. banks failed at a higher rate than small banks and that in later years were no less likely to fail than small banks. Boyd et al. (2006) propose alternative theoretical

models that make contrary predictions about the relationship between bank market concentration and the probability of bank failures. To choose between their models, Boyd et al. examine both a cross-section of about 2,500 U.S. banks in 2003 and a panel of about 2,600 banks in 134 developing nations between 1993 and 2004. They conclude that the failure probability for banks is positively related to concentration. In addition, De Nicolò (2001) has examined U.S. banking data from the 1988–1998 interval, and he concludes that evidence derived from this period implies that the probability of failure increased with bank size.

De Nicolò and Loukoianova (2007) develop a theoretical model that suggests relationships among failure risk, ownership, costs of screening and bankruptcy, and market structure. Tests they conduct using banking data for 133 developing nations for the interval from 1993 to 2004 also yield a positive relationship between bank concentration and risk of failure. Agoraki et al. (2008) reach a similar conclusion based on a study of 546 central European banks over the 1994–2005 period. Furthermore, Carlson and Mitchener (2006) conclude from study of U.S. bank branching data from the 1920s to 1930s that a proliferation of branches increases competition and results in more exit of weak banking institutions, which on net tends to strengthen the overall stability of the banking system.

In addition, Schaeck et al. (2006) apply the Panzar–Rosse (1987) measure of competition to banking industries in 38 nations over the interval spanning 1980–2003 and conclude that more competitive banking systems are less prone to crises. Separately, Schaeck and Čihák (2008) examine 1995–2005 data covering more than 8,900 U.S. banks and more than 3,600 banks in 10 European nations. Schaenk and Čihák utilize a competition measure developed by Boone (2001), which measures the elasticity of bank profits with respect to marginal cost, and they likewise find results consistent with greater stability with increased competition, a result they attribute to increased profit efficiency of more competitive banks.

Beck (2008), Carletti (2008), and Carletti et al. (2002) provide reviews of competition-fragility versus competition-stability hypotheses and a survey of the empirical evidence. Consistent with the above discussion, their judgments are that the evidence at best is mixed. There currently is no firm support for the traditional presumption that banks with greater market power are more stable institutions.

Deposit Insurance, "Too Big to Fail" Doctrine, Basel I, and Basel II

Significantly complicating factors in any assessment of the relationships among bank market competition, concentration, and stability are the structure of deposit insurance and the policy stance known as the too-big-to-fail doctrine. Although Angkinand (2009) provides evidence that nations with deposit insurance systems experience smaller output losses as a result of financial crises, comprehensive deposit insurance distorts choices of both banks and their depositors. Furthermore, enforcement of a too-big-to-fail doctrine almost certainly likewise alters the industry's competitive environment.

The Distorting Effects of Deposit Insurance

In the United States, the Banking Act of 1933, otherwise known as the Glass-Steagall Act, established the Federal Deposit Insurance Corporation (FDIC) and charged this agency with supervising the nation's taxpayer-guaranteed deposit insurance system for commercial banks and savings institutions. After study of U.S. bank failures between 1864 and 1934, the FDIC found that the average cost of such failures per dollar of deposits (up to a limit of $5,000) was about 0.25 percent, implying that this would be an appropriate per-dollar insurance premium to charge banks for deposit insurance. The first head of the FDIC, Leo Crowley, convinced Congress that this premium was too high to make bankers enthusiastic about the new federal system. Crowley also suggested that federal banking regulations imposed as part of the 1933 legislation would reduce bank failure rates relative to historical levels. Consequently, the U.S. Congress specified an initial deposit insurance premium of 0.083 percent. Another 54 years would pass before the FDIC would significantly adjust this premium rate, despite gradual increases in coverage from a $5,000 limit per insured deposit to today's current limit of $100,000 (temporarily raised in 2008 to $250,000).

To avoid creating an adverse selection problem, the U.S. government required all chartered banks to participate in the FDIC's deposit insurance program. Nevertheless, by its nature deposit insurance generated a moral hazard problem: Its existence can lead bank managers to make riskier choices than they might otherwise have made. Hovakimian et al. (2003) show that in nations with strong contractual property rights, market discipline is most readily retained in national deposit insurance systems that include combinations of loss-sharing rules, risk-sensitive insurance premiums, and coverage limits. Absent such features, however, deposit insurance expands risk-shifting incentives. See Demirgüç-Kunt et al. (2008) for a full overview of these and other issues.

A possible solution to the risk-shifting incentives generated by the moral hazard problem of deposit insurance is bank regulation and supervision. By conducting periodic examinations of insured institutions and by supervising the insured institutions through the issuance and enforcement of rules for prudential management, bank regulators potentially can reduce the scope for widespread moral hazard difficulties. Another possible solution is risk-based deposit insurance premiums. Indeed, in 1993 the FDIC put into place a system of risk-based premiums based on both banks' *CAMELS* ratings and their capitalization, with banks exhibiting higher *CAMELS* ratings and lower capital-asset ratios judged riskier and hence required to pay higher premiums. Under this scheme, the lowest premium rate charged the least-risky banking institutions initially was 0.23 percent, very close to the premium rate that the FDIC had judged to be actuarially appropriate 60 years previously.

Under terms of the 1991 FDIC Improvement Act, however, the FDIC was not permitted to charge *any* deposit insurance premiums to well-capitalized institutions with sufficiently low *CAMELS* ratings if the FDIC's deposit insurance fund surpassed 1.25 percent of all insured deposits, a point that was reached by the late 1990s. Thus, for nearly a decade, virtually no U.S. banks paid deposit insurance

premiums. The Federal Deposit Insurance Reform Act of 2005 raised the fund limit for charging premiums to 1.4 percent of insured deposits, and since the late 2000s most banks have paid deposit insurance premium rates ranging from 0.05 to 0.12 percent. Nevertheless, U.S. deposit insurance was for more than a decade been "risk based" in name only. By and large, deposit insurance actually was little more than a taxpayer guarantee to depositors backed by a meager deposit "insurance" fund.

As a consequence, depositors with covered deposits have had little reason to assess the safety and soundness of banks, and in the absence of regulation bank managers have had an incentive make riskier decisions. Thus, government deposit insurance has had distortionary effects that have provided a significant rationale for regulation and supervision aimed at promoting industry safety and soundness.

The Impact of the Too-Big-to-Fail Doctrine

In 1982, a large U.S. bank named Penn Square failed because declining energy prices had caused the market values of many of the bank's energy-related loans to fall dramatically. Penn Square had close financial dealings with a number of other institutions, including Chicago-based Continental Illinois Bank, which at the time was one of the nation's largest banking institutions. Continental Illinois had purchased over $1 billion of Penn Square's energy loans and soon found itself on the same slippery slope toward bankruptcy. When word of Continental Illinois' problems began to spread, it became the victim of an electronic bank run. Depositors whose account balances exceeded the limit for deposits insurance coverage made wire transfers out of their accounts at the bank, causing it to lose over $10 billion in deposits within a two-month period in the spring of 1984. The bank offered above-market interest rates in an effort to induce individuals and firms to purchase its certificates of deposit, and it sold billions of dollars of its assets, but to little avail. By May of 1984 the FDIC had decided to bail out the Continental Illinois by purchasing over $2 billion in subordinated notes from the bank. In addition, the Federal Reserve Bank of Chicago extended long-term credit to the bank.

At that time, these actions by the FDIC and the Federal Reserve were unprecedented efforts to keep a bank from failing, because they protected uninsured depositors of the bank as well as those whose funds were covered by federal guarantees. In September of 1984 the Office of the Comptroller of the Currency, the chief regulator of national banks, announced to Congress that it had determined that the largest eleven nationally chartered banks in the United States were too big to fail. This too-big-to-fail policy had its intended effect of shoring up public confidence in the nation's banking system, and ultimately it was adopted as well by the other federal commercial banking regulators, the Federal Reserve and the FDIC.

Another byproduct of the too-big-to-fail policy, however, was that it effectively gave federal government guarantees to *all* the deposits of the nation's largest banks, insured or uninsured (see, for instance, Stern and Feldman, 2004). Smaller banks, naturally, felt that this gave the largest banks an unfair advantage in the marketplace. In fact, O'Hara and Shaw (1990) have demonstrated that the comptroller's announcement of the too-big-to-fail policy led to a significant increase in the

stock prices of the largest banks in the nation. Stock prices also increased for the four largest *state-chartered* banks in the country at that time—Bankers Trust, Chemical Bank, Manufacturers Hanover, and J.P. Morgan—that were not even on the comptroller's list. Nevertheless, these banks were mentioned in a *Wall Street Journal* article the day after the comptroller's announcement, hence the reaction of investors that boosted their share prices. Thus, being a too-big-to-fail bank offered and, as documented by Deng et al. (2007) and Ennis and Malek (2005), continues offer to a U.S. bank competitive advantages over smaller institutions. This size advantage was enhanced in 2008 when the U.S. Treasury singled out the nation's largest institutions for particularly large government "capital injections" during the financial crisis brought about by the subprime-loan meltdown.

To be sure, as noted by Strahan (2008), small U.S. banks benefit from various explicit and implicit subsidies offered by the government and the Federal Reserve. Nevertheless, not being too big too fail implies the reverse implication that banks operating at lower scales are "sufficiently small to fail," placing these banks at a competitive disadvantage in debt and equity markets. Based on an event study, Kane (2000) suggests that a number of banks in fact have actively sought to become "too big to discipline adequately" in an effort to reduce their funding costs relative to smaller rivals.

Acharya and Yorulmazer (2007) suggest that there is also a "too-many-to-fail" problem associated with the fact that regulators often seek to merge failing or failed banks into healthier institutions. In the context of two- and multi-bank models, they show that pursuit of such a closure policy encourages banks to herd by lending to similar industries or making similar bets on interest-rate movements. They do so realizing that if all banks experience solvency threats simultaneously, the regulator will feel obliged to bail out them out as a group. Acharya and Yorulmazer conclude that small banks are particularly prone to herding behavior to bring about a too-many-too-fail outcome.

Basel I, Capital Regulation, and the Three Pillars of Basel II

In their quest to promote a safer and sounder banking system, regulators have long sought to determine "appropriate" minimal levels of capital, which some evidence, such as that provided by Kwast and Passmore (2000), suggests are depressed by the subsidization effects of deposit insurance and other aspects of the safety net provided to banks. Between the mid-1950s and the mid-1970s, regulators used a formula intended to compute minimum capital levels for different categories of assets. Various financial innovations rendered these formulas useless, however. Thus, in 1981 U.S. bank regulators implemented explicit capital ratios. Regulators required a 5.5 percent minimum ratio of "primary capital"—common and perpetual preferred stock, surplus, undistributed profits, and capital reserves—to total assets, and they imposed a 6.0 percent minimum ratio of "total capital"—primary capital plus certain subordinated notes and debentures, other preferred stock, and mandatory convertible debt—to assets.

In 1988, in an effort to take into account heterogeneities of risks across different sets of bank assets, bank regulators of various nations agreed to adopt the so-called Basel Accord. Under this agreement, now commonly called Basel I, participating nations imposed both a traditional leverage (asset-to-capital-ratio) requirement and "risk-based" requirements relating measures of bank capital to a "risk-weighted" measure of total assets. By the mid-1990s, regulators had determined that banks had learned how to game the system via regulatory arbitrage activities that undermined the intent of the Basel risk adjustments. Since then, considerable regulatory effort has been expended to develop and gradually implement the so-called Basel II system, which aims at constructing and implementing an overall regulatory framework built around three so-called "pillars": Risk-based capital requirements, supervisory discipline, and increased market discipline (see, for instance, Tarullo, 2008). Capital regulation is clearly the central pillar, however, with small banks facing a much more complex risk-based capital-requirement system than under the Basel I system and large banks required to implement an even more sophisticated, internal-ratings-based (IRB) system for capital tabulation.

The Basel II framework was originally scheduled for full implementation by the mid-2000s. In 2008, the system was slated to go into effect gradually with a target completion date of 2011, but in 2009 full implementation was postponed indefinitely pending resolution of the post-2008 financial crisis. Nevertheless, most of the more than 100 adherents to the Basel I system of capital regulation have already expressed intentions, perhaps explained in part by club theory and individual self-interest (see Pattison, 2006), to adopt the Basel II standards.

Summary: Bank Competition, Stability, and Regulation

- The Diamond–Dybvig model rationalizes the existence of a financial intermediary as a mechanism for implementing a risk-sharing agreement among agents with heterogeneous intertemporal consumption timing. This analytical framework motivates bank runs as breakdowns in risk sharing generated by recognition of an intermediary's low earnings on assets owing to economic shocks, a run on another intermediary, or another external event. Difficulties with the Diamond–Dybvig model include the fact that its optimal risk-sharing arrangement can be replicated by a structured mutual fund, the model's dependence on an environment with isolated agents face lower costs in dealing with an intermediary instead of engaging intertemporal trade, and its treatment of depositors as residual claimants instead of holders of demandable debts.
- A characteristic of banking that arguably makes banks "special" is their role in screening and monitoring loans inherently subject to very idiosyncratic risks, thereby improving loan quality and reducing loss risks and bank failure probabilities. Most efforts to empirically evaluate screening- and monitoring-centered theories of the banking industry provide indirect supporting evidence based on responses of the returns of firms receiving credit from banks and hence presumably benefiting from the screening and monitoring functions of banks. Imbedding

monitoring activities into a basic model of the banking industry suggests that banks' loan-monitoring activities affect market outcomes, which in turn feed back to influence the degree to which banks monitor their loans.

- A longstanding contention in banking literature is that unhindered competition makes the banking industry less stable. One strand of this perspective, an "excessive-competition argument," suggests that banks that pay market deposit rates attract funds from depositors more likely to engage in unanticipated withdrawals that expose banks to greater risk of runs. Alternatively, a "competition-illiquidity argument" proposes that in a perfectly competitive market, a given bank has no incentive for other banks to lend to any given bank experiencing liquidity problems even though the banking system as a whole would be better off if such lending were to occur. Furthermore, a competition-asset risk argument contends that an individual bank is more likely to lend to a prospective borrower that has succeeded in obtaining credit from a rival lender resulting in herding behavior and correlated credit risks with a larger number of banks.
- The arguments favoring a negative relationship between bank competition and stability typically fail to take into account the fact that optimal portfolio-risk choices are sensitive to the nature of bank market rivalry and the extent of bank product differentiation and that asset prices and return distributions also may vary with the degree of competition. Empirical evidence regarding the real-world relationship between competition and stability in banking is decidedly mixed, with various studies suggesting evidence of negative, positive, or even U-shaped relationships between the degree of competition in banking markets and the stability of the banking industry.
- In spite of governmental exposure to moral hazard problems created by underpriced deposit insurance, more than a quarter of a century ago U.s. banking regulators further increased governmental exposure to moral hazard risks by adopting a too-big-to-fail policy. In an effort to rein in such risks, during the 1980s and 1990s regulators implemented the Basel I capital standards. Since the early 2000s, national bank regulatory officials have pursued efforts to employ a Basel II regulatory framework encompassing three pillars: a revised capital adequacy regime, a commonly pursued supervisory process, and greater attention to market discipline as a mechanism for containing risks.

Chapter 7
Capital Regulation, Bank Behavior, and Market Structure

In light of five decades of regulatory effort based at least in part on minimum capital requirements, one might anticipate that a perusal of the academic banking literature would yield considerable agreement that capital requirements are a worthy tool within a bank regulator's arsenal. In fact, the theoretical banking literature is sharply divided concerning the effects of capital requirements on bank behavior and, hence, on the risks faced by individual institutions and the banking system as a whole. Some academic work indicates that capital requirements unambiguously contribute to various possible measures of bank stability. In contrast, other work concludes that if anything, capital requirements make banks riskier institutions than they would be in the absence of such requirements.

Why have economists reached such divergent conclusions about the riskiness or stability effects of capital requirements? Researchers have applied a variety of different approaches to analyzing the effects of capital requirements on bank behavior, so answering this question requires conducting a thorough review of the theoretical literature on bank capital regulation. (For reviews of empirical evidence and of broader implications of capital regulation for economic stability and monetary policy, see Jackson et al., 1999, Wang, 2005, and VanHoose, 2008.) Santos (2001) and Stolz (2002) have provided recent surveys of the literature on the stability implications of capital regulation. In contrast to previous surveys, however, this chapter aims to direct a more critical focus upon the reasons for the literature's conflicting conclusions about capital regulation's effects on bank behavior. Furthermore, the present chapter considers a considerable volume of additional academic research on the effects of capital regulation that has emerged since the late 1990s in conjunction with continuing discussion of the appropriate structure and implementation of the Basel II standards.

The Portfolio Management Perspective on Capital Regulation

If a bank is viewed as primarily a manager of a portfolio of assets, then the fundamental effect of any system of capital requirements that actually bind a bank or, in an uncertain environment, that a bank anticipates could prove binding under

D. Van Hoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-642-02821-2_7, © Springer-Verlag Berlin Heidelberg 2010

some circumstances—is to alter the bank's portfolio leverage (asset-capital) ratio. Naturally, from the point of view of portfolio selection, the result will be a change in the composition of the optimal asset portfolio.

The Bank as a Competitive, Mean-Variance Portfolio Manager Facing Capital-Constrained Asset Portfolios

The three seminal analyses of the portfolio impacts of binding capital requirements are the contributions of Kahane (1977), Koehn and Santomero (1980), and Kim and Santomero (1988). The fundamental points of these papers can be understood in the context of the basic mean-variance portfolio-selection model discussed in Chapter 3. Recall that in this type of framework a representative bank takes asset returns as given. The bank determines its optimal portfolio with an aim to maximize the expected utility derived from end-of-period capital, which in turn depends on the relative risk aversion of the bank's owners.

To contemplate the effects of variation in the bank's capital-asset ratio within a portfolio management context, consider Fig 7.1. Suppose that at a capital-asset ratio denoted k_1, the curve labeled EF_1 is the efficient frontier associated with alternative

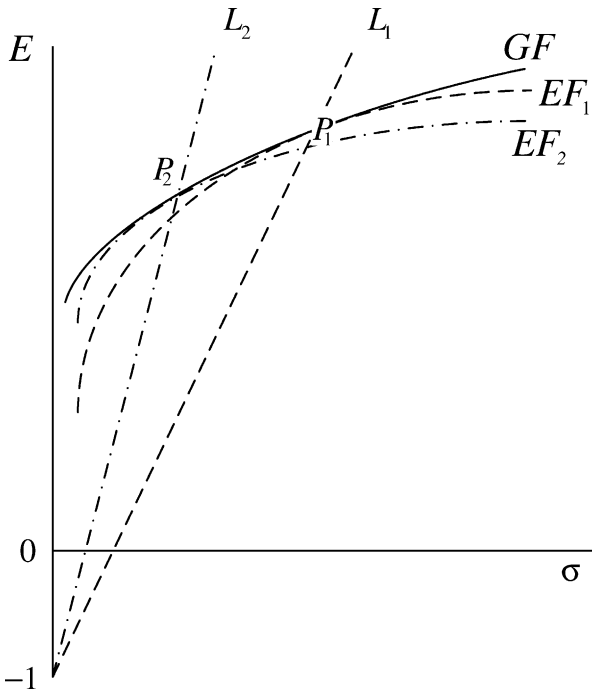

Fig. 7.1 Efficient portfolios and the probabilities of bank insolvency at different capital-asset ratios

values of the expected portfolio return, expressed per unit of equity capital, on the owners' capital investment, E, and of the standard deviation, σ, of the portfolio return. Blair and Heggestad (1978) demonstrate that the slope of the ray L_1 crossing through point P_1 is the square root of the reciprocal of the probability that the mean portfolio return per unit of capital will fall below a value of –1—that is, that the mean end-of-period value of the bank will decline below the level of the bank's capital.

Thus, the steeper ray L_2 corresponding to the portfolio P_2 along the alternative efficient frontier EF_2 is associated with a smaller probability of bank insolvency. The efficient frontier EF_2, in turn, corresponds to a higher capital-asset ratio, denoted k_2; as shown, for instance, by Kim and Santomero (1988), an increase in the capital-asset ratio shifts the efficient frontier leftward and downward.

Figure 7.1 is constructed so that points P_1 and P_2 lie along an envelope of efficient frontiers, including EF_1 and EF_2, corresponding to all feasible capital ratios, such as k_1 and k_2. This envelope is the global frontier GF. A movement downward along GF implies a lower expected return and higher capital-labor ratio and, hence, a less risky portfolio.

To consider the safety-and-soundness implications of capital requirements in this setting, Kahane, Koehn and Santomero, and Kim and Santomero evaluate the effects of a binding leverage constraint on the probability of failure. Suppose, for instance, that the regulator sets a minimum permissible capital-asset ratio $\bar{k}$ orresponding to the probability of failure given by the square root of the reciprocal of the slope of the ray $\overline{L}$ an with the efficient frontier $\overline{EF}$ in Fig. 7.2.

The objective of this minimum capital requirement of $\bar{k}$ is to try to induce banks to select portfolios at (if the capital constraint is binding) or to the left of $\overline{P}$ along the global frontier GF. The difficulty is that not all banks will have risk preferences consistent with such a choice. Only banks that are sufficiently risk averse, with indifference curves at least as steeply shaped as $\bar{I}$, will opt for the portfolios desired by the regulator enforcing the minimum capital ratio. A bank that is willing to accept a greater increase in standard deviation of the portfolio return in return for a larger mean return, such as one with indifference curves as shallow as I_1, will be willing to select a riskier portfolio P_1 along the capital-constrained efficient frontier $\overline{EF}$. Even though banks such as these satisfy the capital regulation, they fail to meet the desired solvency standard.

Kahane, Koehn and Santomero, and Kim and Santomero conclude that a bank that is sufficiently non-risk-averse will respond to a higher capital requirement by choosing a riskier asset mix than before the increase in the leverage ratio, thereby yielding a perverse—from a regulatory perspective—outcome in which the probability of bankruptcy increases. As a consequence, the effect of capital requirements on the overall safety and soundness of the banking system as a whole depends on the distribution of risk aversion across banks. More stringent capital requirements could make some banks safer, some banks riskier, and the banking system as a whole either more or less safe.

Kahane (1977) suggests that capital regulation cannot reduce overall bank portfolio risk unless the asset composition of the bank's portfolio is also subjected to

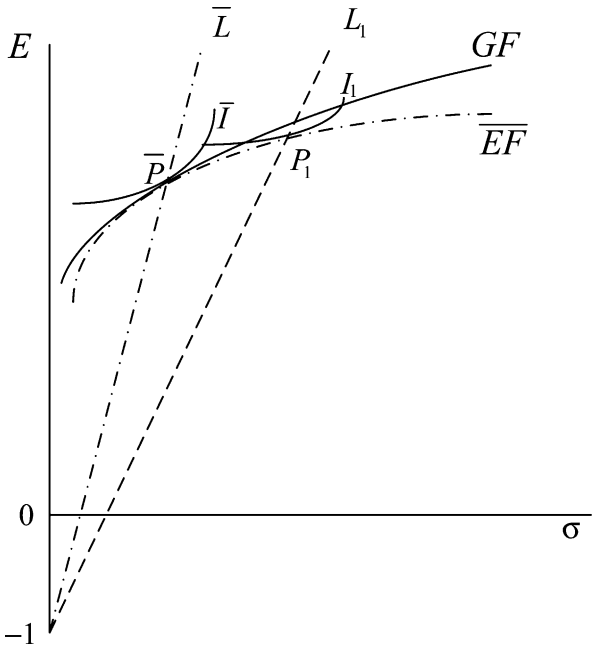

Fig. 7.2 A bank capital requirement and portfolio choice

regulation. One way a regulator might try to address the potential for at least some banks to choose a riskier asset portfolio is to risk-weight banks' assets, as in the Basel I system. Kim and Santomero (1988) extend the portfolio-selection approach to analysis of an asset-risk-weighted system and provide support for this approach, as long as the weights are optimally set (see Bradley et al., 1991, Carey, 2002, Chen et al., 2006, Gjerde and Semmen, 1995, Cordell et al., 1995, and Gordy, 2003 for more detailed discussions of linking capital regulation to risk-weighted assets).

Taking Deposit Insurance Distortions into Account

Kim and Santomero open and close their study by noting that a key rationale for capital requirements is to redress the possibility of greater risk induced by mispricing of deposit insurance. Nevertheless, their analysis does consider the potential behavior-distorting effects of deposit insurance.

Keeley and Furlong (1990) suggest that when a bank's option value of deposit insurance is taken into consideration, it becomes problematic to apply mean-variance analysis in evaluating the effects of capital requirements on the probability of failure. Specifically, when the option value of deposit insurance is recognized, the cost of borrowing is no longer constant, so that the effects of changes in the leverage ratio on risk and return are not linear. Keeley and Furlong argue that the variance

of returns consequently is not an adequate measure of risk, thereby rendering suspect the results obtained in standard portfolio management models. Separately, Furlong and Keeley (1989) incorporate the option value of deposit insurance into a state-preference model of a representative bank with an objective function that is linear in expected return. Their conclusion is that an increase in bank capital is unambiguously associated with a reduction in the level of bank asset risk.

Flannery (1989) independently reaches a similar deduction. He also takes into account the value of the deposit-insurance put option within a model that allows for a broad range of asset choices in the loan portfolio of a risk-neutral bank while simultaneously examining broader balance-sheet impacts of capital regulation. Based on simulations of the model, Flannery concludes that although capital regulation tends to induce a bank to diversify its portfolio less than it would if unregulated, overall risk in a bank's loan portfolio declines.

Gennotte and Pyle (1991) consider a setting in which banks raise a fixed amount of deposits and choose among a set of loan portfolios with differing net present values and risks and extend loans with non-zero net present values. In their framework, the net present value of managers' claims on the bank equals the sum of the call value of the government deposit-insurance subsidy plus the net present value of the bank's assets. Banks invest to the point at which the subsidy on the marginal dollar offsets the negative present value of the marginal investment. To address the distorting effects of the deposit insurance subsidy, the government imposes a capital constraint in the form of a maximum deposit-to-asset value ratio.

The effects of capital requirements on bank risk taking and scale implied by Gennotte and Pyle's analysis are generally ambiguous. If a bank's marginal costs increase with risk—which occurs when the asset portfolio is a combination of investments in safe and risky assets—then the bank responds by increasing the fraction invested in the risky asset, and its scale decreases. There are two effects of a capital tightening on the probability of bankruptcy: (1) reduced leverage, which reduces the bankruptcy probability and (2) increased asset risk, which increases the bankruptcy probability. Which effect dominates depends on a ratio of the elasticities of the net present value of investments with respect to the mean and variance of the present value. Gennotte and Pyle show that if the ratio of marginal to average costs is constant, so that the elasticity ratio is independent of the level of assets, and if asset returns are lognormally distributed, then asset risk definitely increases with a capital tightening. The probability of bankruptcy initially drops with a tougher capital requirement, but this probability ultimately rises with a further tightening.

Explaining the Mixed Implications of Portfolio Management Models

What accounts for these diverging conclusions about the effects of capital requirements portfolio-managing banks? Rochet (1992) suggests that part of the answer depends on whether one assumes complete or incomplete markets. Rochet considers

a setting in which capital is exogenously fixed. In his theoretical framework, representative banks choose the level of deposits, which are subject to increasing marginal costs, and hence their scale. Banks also select a portfolio from a set of risky and riskless assets. Rochet finds that when a capital requirement is unexpectedly introduced into a complete-markets setting with deposit-insured, value-maximizing banks, there are three possible outcomes: (i) no increase in capital but investment in a combination of one risky asset and the riskless asset, (ii) complete specialization in a single risky asset and just meeting the requirement, or (iii) no increase in capital but investment in a combination of two risky assets. As in earlier work based on mean-variance analysis, the risk of failure is most likely to increase in this setting. Indeed, Rochet concludes that risk-based deposit insurance premiums are likely to prove a more effective instrument for reducing portfolio risk than capital requirements. In contrast, in a setting with incomplete markets, limited liability, and shareholder-utility-maximizing banks, results are similar to those obtained by Keeley and Furlong (1990) and Furlong and Keeley (1989): Capital regulation can potentially reduce asset risk. Nevertheless, this condition follows only if risk weights in the required solvency ratio are proportional to the systematic risks of assets as measured by their market betas.

Recently, Jeitschko and Jeung (2005) have sought to consider the roles of managerial agency problems and higher-risk, higher-return assets in influencing the effects of capital requirements on risk at a portfolio-managing bank. Jeitschko and Jeung examine a framework that allows for asset risk orderings more general than mean-variance. They utilize this framework to evaluate responses of bank risk to greater capitalization depending on which agent—deposit insurer, shareholders, or managers—dominates bank decision-making. Jeitschko and Jeung assume that exogenous capital and deposits are invested by the representative bank, with deposits paid back without risk and assets potentially shifted from one investment to another. In addition, regulation precludes a bank from investing in negative net-present-value projects, and their model allows for changes only about the center of the distribution of asset returns that result in a uniform shrinking or stretching of the distribution.

Jeitschko and Jeung find that a bank receiving deposit insurance subsidies may choose a dominated risky asset rated either according to a mean-variance ordering—thereby implying a second-order stochastic dominance relationship—or according to a mean-preserving spread. The expected value of bank equity equals the net return on investment plus the option value of deposit insurance, which is the expected cost of providing the insurance to the bank. Ultimately, the implications of capital regulation of the Jeitschko–Jeung portfolio-management model of bank decision-making depend on which agent dominates in the portfolio decision. If the deposit insurer has the power, say through regulation, to determine the bank's portfolio, its goal is to choose a risk factor that minimizes the option value of deposit insurance. If shareholders dominate, the goal is to choose a risk factor that maximizes the expected value of bank equity. Under strict mean-variance ordering (as in Keeley and Furlong, 1990 and Furlong and Keeley, 1989), greater capital leads to a lower risk factor choice. In the presence of higher-return, higher-risk assets, this result does not necessarily follow, however. The goal of bank managers is to choose the

risk factor that maximizes the expected value of the private benefit of control. The preferred asset risk that results is potentially increasing with greater capitalization, depending on parameter values.

Thus, Rochet and Jeitshko and Jeung have identified several factors that explain the divergent implications of portfolio-management models for the responsiveness of bank portfolio risk to capital regulation. Results depend on whether banks are value-maximizing or utility-maximizing firms, whether bank ownership entails limited liability, and whether banks operate in complete or incomplete asset markets. Furthermore, the effects of capital regulation on portfolio decisions and hence on the banking system's safety and soundness ultimately depend on which perspective dominates in the principal-agent interactions among insurers, shareholders, and managers.

Asset-Liability Management under Capital Regulation

Portfolio management models of capital regulation commonly abstract from the liability side of the bank's balance sheet. One exception is Homölle (2004). In the context of a model developed from state-preference theory, she confirms Furlong and Keeley's (1989) conclusion that a bank capital requirement reduces bank risk in a setting in which a bank issues only insured deposits and has a fixed level of equity on its balance sheet. The reason, she demonstrates, is that a bank must respond to binding capital regulation by reducing its assets and reducing insured deposits commensurately in order to reduce its leverage ratio.

Homölle shows, however, that when a bank is also permitted to issue subordinated, uninsured debts, the resulting broadening of the range of responses across both sides of the bank's balance sheets results in an ambiguous impact on asset risk. In this situation, the bank can respond to a binding capital requirement by altering its issuance of subordinated debt or equity to satisfy the regulation. Homölle finds that when a bank's equity is variable the reaction to a higher capital requirement does not depend on the deposit insurance premium. Whether or not asset risk increases as a consequence depends on how much the bank adjusts its insured deposits in relation to subordinated debt and equity.

Thus, another key factor influencing the predicted effects of capital regulation in portfolio management models is the degree of flexibility that the bank possesses on *both* sides of its balance sheet. This conclusion suggests that failing to account for joint asset-liability decision-making is required to fully assess the effects of regulatory capital restrictions.

An Incentive-Based Perspective on Capital Regulation

Another consideration missing from many portfolio management models of bank capital regulation is an analysis of explicit and implicit costs and benefits that banks face when adjusting to legal restrictions. Milne (2002) suggests that analysis of the

effects of capital requirements on bank asset risk in portfolio management models is flawed because many are insufficient in "treat[ing] banks as forward looking optimizers balancing the benefits of their lending decisions against the costs of a regulatory breach." A number of researchers have given attention to the manner in which a bank balances costs and benefits across its balance sheet when subjected to capital regulation.

Incentives and Capital Requirements

Modern models of constrained bank balance-sheet adjustments to capital requirements focus on the effects of risk-based capital regulation. Several approaches also seek take into account dynamic elements of the problem that banks confront.

Perfect Competition Models of Bank Capital Regulation

Much recent work evaluating incentive effects of bank capital regulation assume perfectly elastic supplies of assets and demands for liabilities on the part of the public. Consider, for instance, the analysis provided by Estrella (2004a), who develops a theoretical framework that, in contrast to portfolio-management models, allows an individual bank to adjust its liability structure. In Estrella's model, a bank facing elements of all three pillars of Basel II makes "staged" decisions. In the first stage, the bank must meet a minimum risk-based capital requirement. In the second stage, the bank must raise funds in the debt market to acquire one of two risky investment assets. In the third stage, the bank obtains a signal about its performance that it may or may not fully pass along to a regulator. Market constraints limit the amount of debt that the bank can issue.

Estrella's model yields equilibrium outcomes implying three possible types of banks: a bank that cannot raise sufficient funds to invest in an asset and closes down, a bank that meets the capital constraint and must issue debt, and a bank that has sufficient capital to invest without issuing debt and that by assumption invests "excess capital" in a riskless asset. A bank regulator's problem is to try to induce a bank to make choices consistent with its own objectives. Higher capital pushes a bank's preferred outcomes closer to those of the regulator but fails to bring them into complete alignment. This gap can be further narrowed by regulatory effort focused on less-capitalized banks. Market discipline also helps close the gap, but it cannot guarantee the first-best result from the regulator's point of view. Estrella suggests that regulatory precommitment to an ex post penalty theoretically could exactly close the gap. He notes, however, that this approach may or may not be practical.

Kopecky and VanHoose (2004a, b) analyze banking systems composed of representative perfectly competitive banks that either are or are not bound by capital requirements. Whereas Kopecky and VanHoose (2004a) contemplates the impacts of binding capital regulation on bank credit supply holding rates of return constant, Kopecky and VanHoose (2004a) allow for adjustments in market interest rates

and add bank equity to the bank's balance-sheet constraint in the basic model of perfectly competitive banking discussed in Chapter 3. Thus, the balance-sheet constraint is $L^i + S^i = (1 - q)D^i + N^i + E^i$ to where E^i is bank equity. Furthermore, the bank confronts a capital-requirement constraint. Under non-risk-based capital requirements, the relevant constraint can be expressed as $\Psi[qD^i + L^i + S^i] \leq E^i$ where Ψ is a minimum allowable non-risk-based capital ratio, while under risk-based capital regulation, the constraint is $\mu[w_R qD^i + w_L L^i + w_S S^i] \leq E^i$, where w_R, w_L, and w_S are risk weights assigned by regulators and μ is the risk-based capital ratio. Under Basel I, $w_R = 0$, $w_L = 1$, and $w_S = 0$, so Kopecky and VanHoose focus on analysis of the capital constraint $\mu L^i \leq E^i$, which naturally holds as an equality when the constraint binds. Kopecky and VanHoose contemplate both a short-run horizon in which equity is fixed at a level $\overline{E}^i$ and a long-run setting in which E^i is a choice variable. In the short run, bank i maximizes its single-period economic profit, $\pi^i = r_L L^i + r_S S^i - r_D D^i - r_N N^i - C^i(L^i, S^i, D^i, N^i)$, whereas in the long run the bank maximizes the sum of current one-period profits and the discounted value of anticipated future profits net of a required payment for equity capital. In both cases, Kopecky and VanHoose obtain explicit solutions for the bank's balance-sheet choices by assuming that $C^i(L^i, S^i, D^i, N^i)$ is quadratic in its arguments.

The essential conclusions forthcoming from the Kopecky–VanHoose model can be depicted diagrammatically in Figs. 7.3 and 7.4, which display market diagrams for representatively capital-constrained banks. Panel (a) of Fig. 7.3 shows that in the short run, with bank equity fixed, a binding risk-based capital constraint effectively acts as a quota on bank lending, so loans decline to $\overline{L}$, and the market loan rate rises above the level in an unconstrained equilibrium, at $\overline{r}_L$. Panel (b) exhibits the effects of an increase in the level of the risk-free rate that translates into higher interest rates

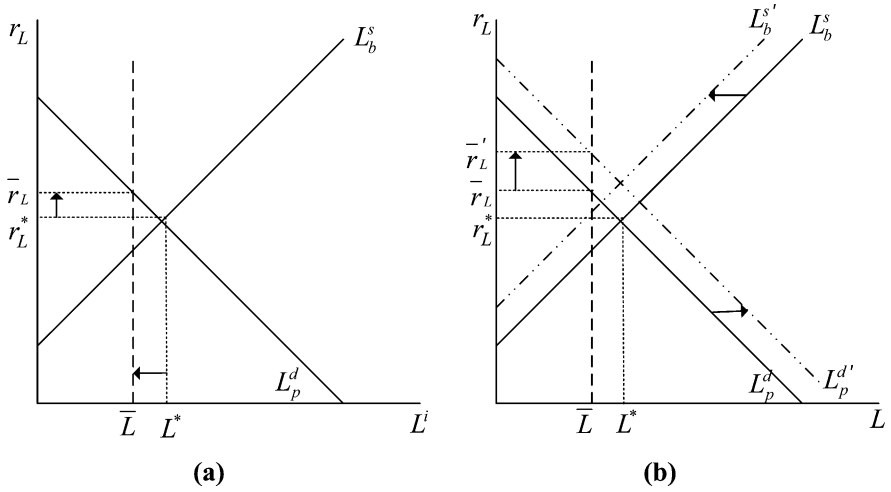

Fig. 7.3 Short-run implications of a binding risk-based capital requirement with representative perfectly competitive banks

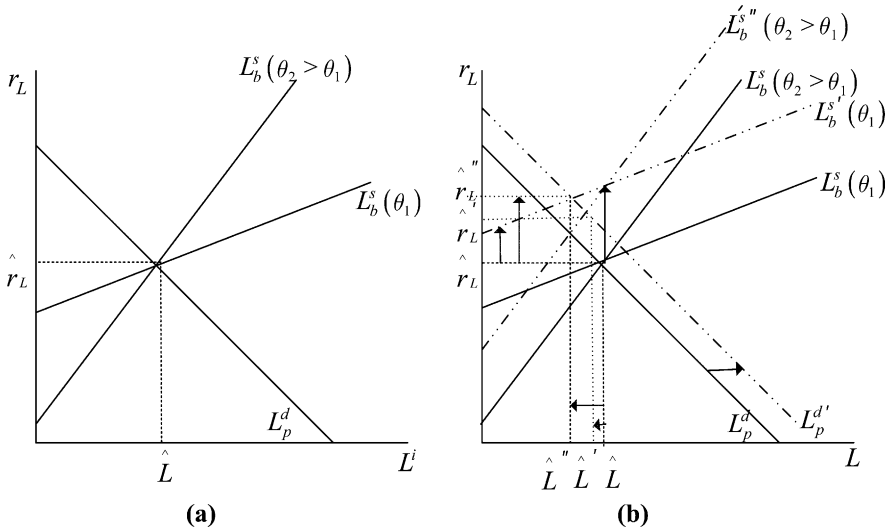

Fig. 7.4 Long-run effects of a binding risk-based capital ratio with representative perfectly competitive banks

on securities and nondeposit liabilities. In an unconstrained equilibrium, the rise in these interest rates raises the total marginal cost of lending and hence causes loan supply to decline; in addition, the public responds by shifting away from securities in favor of bank loans, so loan demand increases. In an unconstrained equilibrium, lending falls in response, and the equilibrium loan rate rises, but in the constrained case the only impact of the higher level of market rates is an increase in the loan rate, from $\overline{r}_L$ to $\overline{r}_L'$.

Panel (a) of Fig. 7.4 shows that in the long run, when equity is variable, banks constrained by capital regulation must raise additional higher-marginal-cost equity to support given volumes of lending, so their total marginal costs of lending are steeper with higher required capital ratios. Consequently, an increase in the required capital ratio steepens the market loan supply curve. Panel (b) displays the effects of an increase in the risk-free rate that boosts interest rates on securities on nondeposit liabilities. Again, the rise in other interest rates generates a decrease in banks' loan supply and a rise in the public's loan demand. In the long run, lending falls, and the equilibrium loan rate rises, with these impacts accentuated under the higher regulatory capital ratio, $\theta_2 > \theta_1$.

Blum (1999) offers a two-period approach showing that the value of equity in later periods is altered when risk-based capital requirements bind an individual bank. In his framework, banks can choose to hold either riskless or risky asset portfolios. Higher expected profits in the second period induce an unregulated bank to reduce risk in the first period to decrease the probability of failure and thereby improve the likelihood of receiving second-period profits. On the one hand, a tighter risk-based capital requirement imposed in the second period reduces that period's

profits in the case of success and hence induces more risk-taking in the first period. On the other hand, a tighter risk-based capital requirement imposed during either the first or second period induces a decline in the feasible allocation of funds to risky assets, thereby reducing risk. If capital requirements bind in the first period only, the latter effect dominates, and risk falls. If capital regulation constrains the bank only in the second period, the bank's risk increases. Blum finds that if capital requirements bind in both periods, the effect on risk is ambiguous.

Calem and Rob (1999) consider an infinite-horizon model in which a representative bank can expand equity only through retained earnings. Bank size is fixed at a normalized value of unity, so their analysis artificially abstracts from scale effects of capital regulation. Deposits are also fixed but costly. A bank can choose portfolio shares of a risky asset and a safe asset. It faces a capital surcharge if its capital is below a minimum standard amount. Both the asset portfolio and the capital position vary over time as a result of past choices and the realization of past risky investments. Calem and Rob trace through the effects of building up a capital cushion, which can ultimately lead a bank to take on more risk in the face of capital requirements, even though over a potentially large range of capital levels portfolio risk declines with increased capital.

Calem and Robb consider calibrations of bank responses to capital regulation under various initial conditions using 1984–1993 data, which yields a cross section of calibrated investment choices of banks with different capital positions. This approach permits some consideration of heterogeneities in banks' responses but does not allow these heterogeneities to affect any market outcomes. It indicates that a U-shaped relationship can arise between capital position and risk-taking. Undercapitalized banks take on maximum risk. At first, risk declines as capital increases, but risk can potentially increase as capital increases beyond a certain point. Also, premium surcharges on undercapitalized banks generate a widening of the range of capital levels over which undercapitalized banks take on maximum risk. Finally, market pricing of uninsured liabilities ("market discipline") may deter excessive risk-taking by undercapitalized banks, but only if risk is priced *ex ante* in response to changes in banks' portfolio risks.

Milne (2002) contends that the incentive effects generated by capital regulation have not received sufficient attention, because most of the literature assumes that capital requirements are viewed as ex ante binding constraints. He argues that the main effects of capital regulation operate through banks' efforts to avoid ex post penalties imposed by regulators if violations of capital adequacy standards take place. This perspective suggests seeking to reduce banks' risk-taking behavior by toughening regulatory penalties rather than assessing more stringent or more requirements tied to asset risks.

Along these lines, Blum (2008) focuses on the truth-telling properties of the internal-ratings-based (IRB) procedure specified under the Basel II system of capital regulation when supervisors possessing potentially differing capabilities to enforce penalties imperfectly detect dishonest revelations under an IRB-style approach. He concludes that inclusion of a leverage-ratio requirement alongside an IRB-style structure of capital regulation—slated to be required in the United States but

presently not in Europe—ensures that banks will voluntarily reveal their risks. In Blum's model, this conclusion is driven by two factors. First, requiring banks to hold a higher level of capital increases downside risks, which reduces the put option value of limited liability. Second, increasing the leverage ratio boosts the anticipated sanction that banks confront. Together, these effects of a higher leverage ratio increase the incentive for banks to truthfully reveal their risks under an IRB-style system.

Marshall and Prescott (2001, 2006) have explored how regulators might use state-contingent pecuniary penalties to induce banks to control risks. Marshall and Prescott (2001) consider the use of ex post fines to partially substitute for capital regulation in the context of a theoretical framework in which a bank chooses its capital, a level of costly screening that determines its average portfolio return, and its portfolio risk. The latter two choices are the bank's private information, and because the bank's deposits are insured it selects a level of screening that is lower than the social optimum and portfolio risk that exceeds the social optimum. In the environment considered by Marshall and Prescott, *ex ante* capital requirements unambiguously reduce the bank's risk level. Imposing a schedule of ex post penalties—fines on banks with high returns generated by selecting a portfolio with considerable risks and on banks with low profits revealing a low level of screening—enables a regulator to reduce capital requirements while deterring banks from expanding their portfolio risks.

Marshall and Prescott (2006) elaborate on this analysis by adding unobservable heterogeneity in the quality of banks' screening abilities. They find that taking into account this additional form of private information indicates that capital regulation fails to provide a means of distinguishing between high- and low-quality screeners. In addition, this additional source of heterogeneity does not to affect their prior conclusion that particularly high returns should be penalized as indicative of higher-risk portfolios. The main alteration resulting from taking into account bank heterogeneity in quality of screening is that charging a fine on good-quality screeners with lower returns induces low-quality screeners to reveal this private information.

Monopolistic Competition Models of Capital Regulation

A number of researchers have contemplated the effects of capital regulation on imperfectly competitive banks. Hellmann, Murdock, and Stiglitz (2000) consider a setting with monopolistically competitive deposit markets, in which the perfectly competitive limit of infinite elasticity of deposit supply yields a deposit rate equal to the riskless rate and a bank earns no economic profits. An individual bank can choose between holding a riskless ("prudent") asset with known return or a risky ("gambling") asset with a random return, which in a successful state is higher than the riskless asset's return. The opportunity cost of capital is higher than the rate of return on the riskless asset. If the bank chooses the risky asset and earns the lower return, it fails. An equilibrium deposit rate above a critical threshold yields each representative bank greater expected returns than investing in the riskless asset. Hence,

in an unregulated deposit-market equilibrium, banks hold undiversified portfolios consisting only of risky assets.

Hellmann et al. find that imposing sufficiently stringent capital regulation reduces banks' incentive to mobilize as many deposits to fund risky assets and hence is Pareto improving, although when returns on assets and bank capital are endogenous, banks' incentives to take on greater asset risk increase when capital requirements are toughened. Hellmann et al. conclude, however, that combining a deposit rate ceiling (for an earlier examination of issues relating to deposit rate ceilings, see Lam and Chen, 1985) with capital regulation can unambiguously induce all banks to reduce investment in risky assets. Thus, Hellmann et al. provide a rationale for coupling capital regulation with deposit rate ceilings, consistent with the perspective, discussed in Chapter 6, that unrestrained competition in bank deposit markets makes banks less safe.

Repullo (2004) build on Hellmann et al.'s approach by applying a Salop-style model of spatial monopolistic competition to the market for deposits in which, as discussed in Chapter 3, banks are located equidistant along a unit circle. Overlapping-generations depositors face costs of moving between banks, which results in deposit demand that depends on the number of banks, the spread between banks' deposit rates, and the mobility cost. In Repullo's model, representative banks facing a regulatory capital constraint choose between riskless and risky assets with exogenous returns. As in Hellmann et al., Repullo finds that either imposing capital requirements or a deposit rate ceiling raises the bank's expected operating margin, which gives it greater incentive to invest in the prudent asset. In addition, however, risk-based capital requirements better constrain banks to prudent assets with less harm to depositors' welfare than deposit rate ceilings.

Repullo and Suarez (2004) provide an alternative analysis of capital regulation in a Hellmann et al.-style setting with imperfect competition in deposit markets and moral hazard in lending, but in the context of overlapping generations of depositors. As in Repullo's analysis, their basic conclusion is that capital requirements are always effective in controlling bank risk-taking incentives. They find that deposit rate ceilings can potentially expand the range of equilibria in which banks choose high-risk portfolios. Repullo and Suarez's analysis also suggests potential differences in effectiveness of risk- versus non-risk-based capital regulation. Furthermore, it indicates that regulations intended to discourage banks from selecting high-risk portfolios are more likely to be successful when banks' market power is greatest, so that banks have less incentive to gamble.

There are good reasons to wonder whether the idea of coupling capital regulation with deposit ceilings would actually achieve real-world welfare improvements forthcoming in the model banking systems proposed by Hellmann et al., Repullo, and Repullo and Suarez. In these papers, asset returns are unaffected by capital regulation, but in reality asset returns—for instance, market loan rates—should respond to market forces that depend in part on the decisions that banks make when confronting capital requirements. In addition, the presumed structures of deposit markets in both studies impose an implicit assumption that there are no close substitutes for bank deposits. Thus, both analyses hinge on the hypothesis that deposit rate ceilings can

be imposed without inducing disintermediation, even though the U.S. experience with such ceilings during the 1960s and 1970s indicates that this hypothesis receives little empirical support.

Demandable Debt, Bank Risks, and Capital Regulation

As discussed in previous chapters, a large part of the banking literature emphasizes the various types of moral hazard problems confronted by banks. Thus, it is not surprising that this topic receives considerable attention in work analyzing bank capital regulation. Indeed, theories of bank capital increasingly focus on how agency and moral hazard problems influence a bank's capital choice and more broadly influence its entire balance sheet.

Obviously, when banks are constrained by capital requirements, regulator-determined capital and market capital are not equal. Berger et al. (1995) provide a helpful discussion of various rationales for regulator-determined versus market capital ratios in light of the Modigliani-Miller theorem. In the absence of capital regulation, they argue, key motives for bank equity relate in part to a trade-off between providing a cushion by issuing equity and gaining tax advantages from debt issuance. Furthermore, Berger et al. suggest that in the presence of asymmetric information, (i) higher bank equity signals to capital markets that bank insiders consider their assets to be riskier; (ii) agency conflicts may develop among bank shareholders, managers, and creditors, such that "[h]igher capital avoids expropriation problems between shareholders and creditors but aggravates conflicts of interest between shareholders and managers" (p. 399). Nevertheless, government safety-net guarantees reduce the incentive to issue equity shares, causing market capital levels to be artificially reduced. Hence, banks face a number of agency problems and associated moral hazard risks that impinge on the capital decision without and with capital regulation.

Capital Regulation and Fragile Deposits

One perspective emphasizes the role of moral hazard problems on the liability side of banks' balance sheets. Diamond and Rajan (2000, 2001) are important contributions in this vein. Building on earlier contributions by Diamond and Dybvig (1983) and Calomiris and Kahn (1991), Diamond and Rajan (2001) develop a theory of banking in which fragility to runs commits banks to creating liquidity. Increased liquidity enables depositors to withdraw upon demand while buffering borrowers by permitting banks to continue to extend credit. In their view, financial fragility incorporated through demand deposits allows a bank to fund itself at low cost, disciplines bank rent extraction, and enables the bank to provide liquidity to both depositors and borrowers.

Diamond and Rajan (2000) extend this approach to develop a theory of bank capital. They trace decision trees for entrepreneur/borrowers, banks/lenders, capital/debt

holders, and depositors, and work out conditions for various choices by all agents. A key implication of their emphasis on financial fragility on the liability side of the bank's balance sheet is that the demandable nature of deposits is crucial to explaining a bank's optimal capital choices. The potential for a deposit run serves to discipline the bank, and the main role of capital is to give a bank a party with which it can negotiate when a bad outcome occurs—essentially a rationale for the old "capital-as-cushion" idea. A clear-cut implication is that a bank's leverage ratio should increase when the underlying liquidity of projects increases—which they note is consistent with the upward trend in leverage ratios over the past decades. The model yields separating equilibria in which, depending on their capital structures, banks seek out particular entrepreneurs to whom to lend.

Within the Diamond–Rajan (2001) framework, the short-run effects of binding capital requirements are a credit crunch for cash-poor borrowers and smaller loan repayments for cash-rich borrowers. Thus, capital requirements have redistributional effects among borrowers. Diamond and Rajan also find that capital requirements can actually increase the chance of a run, because the requirements encourage banks to liquidate sooner, effectively reducing the amount they collect and their ability to honor deposit contracts. In the long run, Diamond and Rajan conclude (p. 2455), capital regulation has ". . .subtle effects, affecting the flow of credit and even making the bank riskier." They continue: "These effects emerge only when the capital requirements are seen in the context of the functions the bank performs rather than in isolation." Deposit insurance only complicates outcomes when *all* deposits are insured, in which case, "deposits are essentially no different from capital" (pp. 2455–2456). If some deposits are uninsured, then the basic results about capital in their paper still go through because the uninsured deposits provide a motivation for capital.

Cooper and Ross (2002) draw on the liability-focused approach to evaluate the interplay of deposit insurance and capital regulation in an environment susceptible to liquidity crises and bank runs. In the context of a basic Diamond–Dybvig two-period consumption model with risk-averse depositors and moral hazard, depositor monitoring can induce the bank to adopt the depositors' desired portfolios. Although deposit insurance removes the threat of bank runs, it reduces the incentive for depositors to monitor the bank's asset choices. A sufficiently high capital requirement related to the level of deposits induces the bank to choose safer assets, thereby mitigating the moral hazard problem that depositors face.

Dowd (2000) argues that maintenance of a sufficient capital cushion can in fact fully solve the financial fragility problem emphasized by Diamond and Rajan. Dowd re-evaluates the original Diamond–Dybvig motivation for deposit insurance by adding a third agent who provides capital and charges a liquidity insurance premium (implicit and deducted from the return paid to depositors). The provision of liquidity insurance by a capitalized "bank" prevents a "run" as long as sufficient capital is provided. Thus, in Dowd's view sufficient capital prevents liquidity crises from occurring. A follow-up paper by Marini (2003) builds on Dowd's analysis to argue as well that market-capitalized banks also should not experience insolvency crises.

Taken together, the Dowd and Marini arguments suggest that market-determined levels of bank capital can substitute for both regulatory oversight and a financial safety net.

Moral Hazard, Bank Lending and Monitoring, and Capital Regulation

As discussed in Chapter 6, work by Diamond (1984, 1991) and others has emphasized the key role that banks play in monitoring loans for moral hazard risks on the asset side of their balance sheets. An interesting examination of how capital requirements alter the incentives that banks face in monitoring loan quality is contained in Besanko and Kanatas (1996). Building on Gennotte and Pyle's analysis of bank portfolio management when non-zero net present value investments are considered, Besanko and Kanatas emphasize the joint effects of deposit insurance and capital requirements, with a representative bank able to fund itself at the risk-free rate and make positive net-present-value loans that ensure a positive surplus to shareholders. Besanko and Kanatas analyze a setting with a representative bank and four agents—bank insiders, bank outsiders, depositors, and a bank regulator. Existing "insider" shareholders own part of bank's equity, make loan portfolio decision in the first period from an endowment of risky loan opportunities, and make a decision about disutility-generating effort that increases the probability of loan repayment. Complying with a capital requirement requires raising more equity from new, "outsider" shareholders.

In the analysis of Besanko and Kanatas, requiring a bank to substitute equity for deposit financing through capital regulation cuts into shareholders' surplus. This creates a potential for an agency problem to arise when the bank must issue new equity to meet a capital adequacy requirement. New "outsider" shareholders help compensate existing, "insider" shareholders by paying a market rate of return on equity. Nevertheless, "insiders'" portion of the surplus contingent on loan repayment declines, which induces insiders to put less effort into realizing a positive-net-present-value loan outcome. Thus, faced with dilution of their surplus, inside shareholders have less incentive to monitor loans, so the probability of loss on loans increases. Because participants in the market for bank equity shares realize that insiders will become less productive monitors, the bank's market value drops. A subsidiary implication is that the regulator must engage in costly monitoring to make sure that insiders do not cut back on monitoring. This places a burden on regulators that, together with the drop in banks' market values, give regulators an incentive to ignore the capital requirement.

Campbell, Chan, and Marino (1992) also consider a model in which monitoring for moral hazard risks is an important function of banks. They examine a representative-bank model in which depositors (or agents they hire, such as a deposit insurer) either monitor banks or impose capital requirements. In the Campbell et al.

framework, all agents are risk-neutral, and banks can choose between two investment projects with the same returns in three states but with different probabilities of arising in the three states. Depositors/insurers choose the capital ratio and/or effort devoted to monitoring to obtain signals of the probability differences, but they do so not knowing the difference in underlying probabilities of project returns. A bank, which knows these probability differences, selects a payment promised to depositors that is designed to maximize the return on the bank's portfolio of projects, subject to the capital/monitoring choices of depositors/regulators and subject to a break-even condition in which the gross portfolio payoff is at least equal to the initial capital investment. In this setting, depositors/insurers desire to increase both monitoring and capital requirements when overall portfolio risk increases the incentive. If problems for monitoring become more severe, however, monitoring and capital requirements are substitutes. Then depositors/insurers prefer to increase capital requirements and engage in less monitoring.

Prior to the bank solvency meltdown between 2007 and 2009, most banks had capital positions in excess of those required by regulation. Allen et al. (2008) suggest that the utilization of relatively expensive capital signals depositors that the bank is committed to monitoring its loans to ensure its soundness, thereby assuring depositors of safety and enabling the bank to raise deposit funds more cheaply. In a competitive market, Allen et al. conclude, the precommitment to increased monitoring makes a more capitalized bank to prospective borrowers and improves a bank's credit-market opportunities.

In Santos (1999), two sources of moral hazard exist simultaneously: one involving the bank relative to the deposit insurance provider and another involving the borrower and the bank. In his model, risk-neutral entrepreneurs earn additional rents when projects succeed, and risk-neutral banks face a cost of capital that exceeds the risk-free rate and capital regulation. Payments to the bank from the entrepreneurs are negotiated in advance, and the bank cannot observe the entrepreneurs' effort levels. Depositors are risk-averse, and a deposit insurer charges banks a flat-rate deposit insurance premium that is anticipated to allow the insurer to break even. The benchmark for evaluating outcomes under moral hazard is the first-best equilibrium in which there are no moral hazard problems. Santos applies this setup to an evaluation of first- and second-best contracts without and with sources of moral hazard in representative banks' lending and deposit-insurance relationships and of the manner in which tougher capital requirements affect those contracts. Because capital is more expensive than deposits, banks always choose the minimum capital level specified by regulators and hence are always bound. The optimal contract between a bank and an entrepreneur entails the bank extending a loan *and* having an equity stake in the firm, but the contract is distorted by deposit insurance. Increasing the required capital standard results in a contract adjustment that takes into account both higher costs that would have been incurred in bankruptcy when its leverage was lower and the higher cost of required capital. The result is lower bank risk of insolvency. Hence, the Santos analysis implies that capital regulation unambiguously reduces bank risk.

Capital Regulation and Bank Heterogeneities

Most of the literature examining the role of asymmetric information as a factor influencing the effects of capital regulation on bank decision-making focuses on moral hazard problems faced by representative banks. Nevertheless, some attention has also been given to how adverse selection problems influence the impacts of capital regulation. In addition, recent work has begun exploring how differences among banks can affect market outcomes under capital requirements and investigating how market adjustments in turn impinge on heterogeneous banks' responses to capital regulation.

Adverse Selection and Capital Regulation

Thakor (1996) offers a unique analysis of the effects of capital regulation on bank behavior that focuses attention on implications for bank screening of prospective borrowers in the face of the possibility of adverse selection (banks also monitor borrowers ex post). Thakor considers settings with both a monopoly bank and multiple, representative banks. In the latter case, a Nash equilibrium reflects each bank's strategy to select an interest rate at which it is willing to lend to a screened borrower in order to maximize its profits net of screening costs it incurs. In equilibrium, a loan applicant's probability of being rationed and receiving no credit increases with a higher cost of funds, which can result from a tightened capital adequacy requirement. Hence, a toughened capital requirement results in more credit rationing on the part of banks. Aggregate bank credit declines.

Deposit insurers and bank regulators must also screen applicants for banking licenses in an effort to combat adverse selection problems that they confront. This is the topic of a recent contribution by Morrison and White (2005), which explores how regulatory screening ability influences the optimal setting of capital requirements. Morrison and White build on work by Holmström and Tirole (1997) to examine an economy inhabited by diverse agents, each of whom has a dollar to invest in projects that either succeed or fail. A fraction of agents with the capability to monitor projects ("sound" agents) are able to increase the probability of a higher return by so doing. Agents can alternatively deposit their endowments with other agents who act as "banks" and pay a fee to the bank for the opportunity to receive the return on a successful project, though neither banks nor depositors earn any gross return when projects fail. Sound agents can earn higher returns, so welfare is maximized when all funds are handled by sound agents. Monitoring must be incentive compatible for sound agents, requiring earnings derived from deposits—the differential between the banks' return on investment and the lower rate paid to depositors—to be high enough to cover their monitoring costs.

In this environment, Morrison and White find that an unregulated (no entry restrictions) banking system can be efficient (only sound agents open banks) only if the cost of monitoring is sufficiently small. Thus, if monitoring costs are above a critical level consistent with efficiency, regulators may be able to increase efficiency

by imposing capital adequacy requirements, restricting entry via screening and licensing, and/or auditing banks. Tight regulatory policy that includes relatively high capital requirements is more likely to improve efficiency if regulatory screening ability is low. If regulators have high screening ability, then looser regulatory policy with lower capital requirements can improve efficiency. In the event of a banking crisis arising from a shift from optimistic to pessimistic expectations (which only can occur if regulators are not of very strong ability), weaker-ability regulators will tend to tighten capital requirements, and stronger-ability regulators will tend to loosen capital requirements.

Capital Requirements, Heterogeneous Banks, and Industry Structure

Morrison and White's (2005) analysis of adverse selection effects of capital requirements requires consideration of heterogeneous banks. In the literature on effects of capital regulation on bank behavior, contemplating diverse banks within the same model is a novel idea. Studies have almost exclusively focused attention either on responses by a single, presumably "representative" bank or on a banking system made up of identical banks. By their nature, however, representative-bank models fail to capture feedback effects between bank-level choices and market-level outcomes. Such theoretical frameworks bear little resemblance to world banking systems composed of institutions displaying diverse management capabilities and utilizing heterogeneous levels of technological sophistication. Furthermore, because representative-bank models yield the conclusion that all institutions are either bound or not bound by capital regulation, these models are always inconsistent the real-world observation that only a small fraction of a banking system is typically constrained by capital requirements.

One recent effort to allow for diversity in responses to capital regulation is provided by Almazan (2002), who considers an environment in which banks balance capital versus monitoring expertise, which is proxied by selected distance of bank from borrower. If the bank locates itself farther from a borrower, then it must commit more capital, so that "capital is a tool that allows a bank to offer lower loan rates without affecting its incentives to monitor" (p. 89). There are three types of agents—investors who own uninformed capital, entrepreneurs who lack capital but are endowed with projects yielding risky payoffs, and *two* banks endowed with capital and possessing monitoring technology. If a bank incurs a cost, it can reduce the borrower's private benefit from pursuing a higher-risk project. In one settings, location (and hence expertise) is fixed. If there is just one bank, then the distance of the marginal project from the bank decreases (that is, bank's expertise increases) with increases in monitoring cost, the riskless interest rate and the size of entrepreneur's private benefit, or with decreases in the net present value of the entrepreneur's project and the bank's capital. In another setting there are two banks, one of which is endowed with more capital than the other, and location (expertise) is endogenously determined.

Four possible equilibrium outcomes arise in Almazan's framework: scarce capital with no interactions between banks, plentiful capital with no rents to banks and hence no monitoring, a case in which only one bank monitors, and a case in which both banks monitor. Almazan focuses on the last equilibrium, in which greater overall capital leads to higher market share for a bank endowed with more capital and in which a rise in the riskless interest rate favors this bank also, with the implication that contractionary monetary policy hurts smaller banks more than larger banks. The key finding in this setting is that the bank endowed with less capital prefers more "separation" from the other bank. Hence, the low-capital bank prefers to specialize in a niche involving less monitoring.

Acharya (2009) considers a model in which limited liability and cross-bank failure externalities induce banks to herd into risk-correlated investments. In his model, because capital-requirement policies are directed at individual banks instead of the banking industry as a whole, the application of such policies can actually boost overall systemic risk in some circumstances. Acharya contends, therefore, that effective risk-containment policies must be applied to banks collectively rather than individually.

Kopecky and VanHoose (2006) examine the effects of capital regulation in the heterogeneous-bank framework discussed in Chapter 6, in which the fraction Ω of competitive banks monitors loans for moral hazard risks and the fraction $1 - \Omega$ does not, with all banks in the industry facing the market loan demand function, $L^d = l_0 - l_1 r_L$. Under capital regulation, an individual bank i's profit function is

$$\hat{\pi}^i = \hat{R}_L^i L^i + r_S S^i - r_D D^i - \frac{\alpha}{2}(L^i)^2 - \frac{\beta}{2}(S^i)^2 - \frac{\theta}{2}(D^i)^2 - \left(\frac{i}{1-i}\right)\frac{c}{2}(L^i)^2,$$

where $\left(\frac{i}{1-i}\right)\frac{c}{2}(L^i)^2$ is bank i's monitoring cost and where $\hat{R}_L^M = \eta r_L$ for a bank that monitors and $\hat{R}_L^{NM} = (\eta - \delta) r_L$ for a bank that does not. A bank constrained by a risk-based capital requirement maximizes its profits subject to its balance-sheet constraint, $L^i + S^i = (1-q)D^i$, and, if capital requirements are binding, to the binding capital constraint $L^i = \hat{L}^i = \frac{\overline{E}}{\mu}$, where μ again denotes the capital requirement ratio.

In this setting, imposing capital requirements alters the terms of the cost-benefit analysis of constrained banks contemplating whether or not to monitor. Lending responses of capital-constrained banks influence the market loan rate and thereby affect the cost-benefit analysis of *unconstrained* banks as well. Kopecky and VanHoose first consider the case in which the entire banking industry is bound by capital regulation, in which case both monitoring and nonmonitoring banks are constrained to the same balance-sheet decisions. Aggregating this constraint across all banks and combining with the loan demand function yields the loan market equilibrium (LME) locus $\hat{r}_L = \frac{l_0 - (\overline{E}/\mu)}{l_1}$, which is the horizontal schedule depicted in panel (a) of Fig. 7.5. The loan rate that yields zero economic profits for both monitoring and nonmonitoring banks turns out to be $\hat{r}_L = \frac{1}{2}\left(\frac{\Omega}{1-\Omega}\right)\left(\frac{c}{\delta}\right)\left(\frac{\overline{E}}{\mu}\right)$, yielding the equal profits (EP) locus also depicted in panel (a). In contrast with the case of the case of unregulated banking industry considered in Chapter 6, under capital

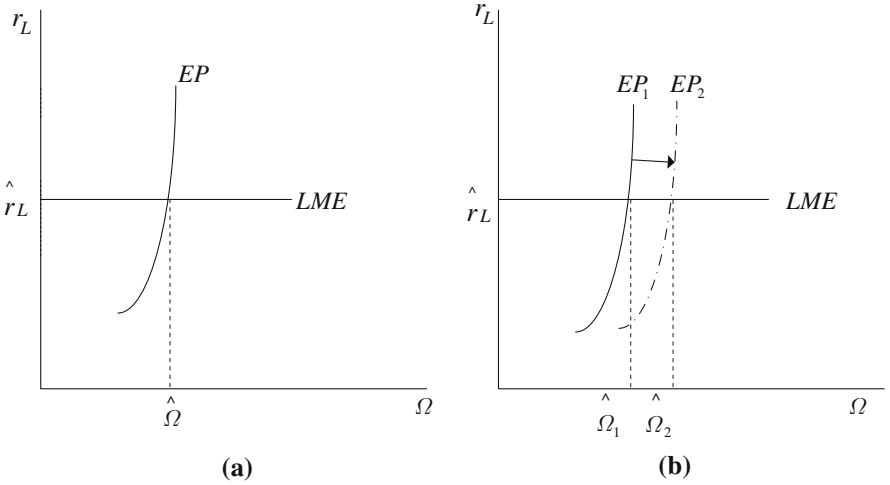

Fig. 7.5 Effects of a binding risk-based capital ratio with heterogeneous competitive banks

regulation the EP locus slopes upward, because as the loan rate rises a bank on the margin regarding a choice whether or not to monitor experiences higher additional profit if it monitors, requiring a higher value of Ω to boosts monitoring costs and re-equalize profits with those of nonmonitoring banks. Kopecky and VanHoose utilize a calibrated simulation to show that the initial equilibrium fraction of banks that monitor their loans, depicted in panel (a) as $\hat{\Omega}$, lies below the equilibrium fraction that otherwise would have monitored their loans in the absence of capital regulation (the fraction Ω^* in Fig. 6.1 in Chapter 6). Kopecky and VanHoose conclude, therefore, that other things being equal, imposing capital requirements that bind the entire banking industry reduce the amount of lending, which reduces risk exposures of banks. At the same time, however, all-binding capital regulation generates a reduction in the extent of loan monitoring and hence tends to make the industry less safe. On net, the safety-and-soundness effects of capital requirements are ambiguous.

Once all-binding capital requirements are in place, increasing the required capital ratio μ shifts the EP locus to the right, as shown in panel (b) of Fig. 7.5, because the resulting further reduction in constrained bank lending reduces marginal monitoring costs, thereby giving some banks a greater incentive to monitor their loans. Thus, after a regime of capital standards that binds all banks is in place, boosting the required capital ratio tends to generate an improvement in the safety-and-soundness status of the banking industry.

The other situation that Kopecky and VanHoose contemplate is one in which all nonmonitoring banks are just constrained by capital regulation, so that the subset of monitoring banks that would have extended larger quantities of loans also are constrained. In their calibrated simulation of this situation, Kopecky and VanHoose obtain conclusions qualitatively analogous to those for the case in which all banks are bound by capital requirements. On net, therefore, their conclusion is

that when capital requirements fully or partially constrain the banking system, lending declines, but so does the equilibrium share of banks that optimally choose to monitor their loans for moral hazard risks. As a consequence, aggregate loan quality may improve or worsen.

In contrast to Kopecky and VanHoose (2006), Boot and Marinč (2006) allow for the banking industry's long-run structure to change in response to market entry. Boot and Marinč examine a setting in which diverse banks make costly investments in monitoring technologies. The magnitudes of these investments in turn alter the benefits derived from monitoring. Boot and Marinč consider a capital-regulated environment throughout and hence do not compare their results with those that would follow in a capital-unconstrained banking system. In this capital-constrained setting, banks endowed with "good" or "bad" quality types take into account the behavior of all market rivals when choosing both loan rates to offer to borrowers and investment expenses to incur in monitoring technologies. "Good" banks choose to undertake more monitoring investment than "bad" banks. Higher capital requirements reduce the size of the deposit insurance subsidy received by all banks and thereby give banks a greater incentive to internalize risks. "Good" banks respond by seeking to reduce their risks by investing more in monitoring technologies, which enables them to reduce their per-unit costs and expand their market shares. In contrast, for "bad" banks the resulting changes in competition involving loan rates and monitoring technologies boosts per-unit costs and causes them to lose marginal borrowers. Hence, in the near term tougher capital regulation strengthens high-quality banks at the expense of low-quality banks, which suffer drops in their market values.

In the longer term, one possible implication—one of several interesting predictions—of the Boot–Marinč analysis is a strengthened banking system as capital regulation ultimately weeds out the weakest banks. Boot and Marinč show, however, that at intermediate levels of quality and sufficiently high degrees of competition, a banking system open to entry could experience a reduction in monitoring incentives. Thus, the impact of capital regulation on aggregate loan quality is ambiguous.

There has been relatively little empirical work evaluating potential impacts of capital regulation on the competitive structure of the banking industry. Recent exceptions are Berger (2006), who analyzes the types of loans issued by banks of different sizes and concludes that large banks failing to utilize internal-ratings-based methodologies instead of a standardized approach under Basel II could be placed at a competitive disadvantage, and Lang et al. (2008), who reach the opposite conclusion with respect to large credit card lenders.

Capital Regulation, Credit Shocks, and Procyclicality and Risk

As noted by Bliss and Kaufman (2003), because bank capital regulation impinges on balance-sheet responses of the banking system as a whole, capital requirements potentially can affect the broader economy. During a short-run interval in which

adjusting equity may prove costly, much of the adjustment to a regulatory capital tightening will tend to occur via reductions in lending. Hence, it is possible that regulatory tightening of capital requirements could transmit short-term external shocks to aggregate credit and hence to the economy. In addition, there is a potential for capital regulation to contribute to procyclical variations in total credit that may create procyclical movements in other economic variables.

Does Toughening Capital Requirements Boost Bank Capital Ratios and Create Credit Shocks?

Interestingly, there is not strong evidence that the imposition of capital regulation has contributed significantly either to an increase in actual bank capital ratios or a reduced level of bank risk. Based on estimates derived from value-at-risk models, Hendricks and Hirtle (1997) conclude that capital regulation is likely to boost capital levels only very slightly at most institutions (and possibly reduce capital at some banks). Ashcraft (2001) also finds little evidence that capital regulation during the 1980s materially influenced bank capital ratios. Flannery and Rangan (2004) find some influence of capital regulation on actual bank capital ratios, but they credit greater bank risk aversion and actual risk increases as the main factors accounting for rising U.S. capital ratios up to the late 2000s.

Other authors reach more mixed or even negative conclusions regarding the contribution of capital regulation to bank equity adjustments. Building on the simultaneous-equations estimation approach developed by Shrieves and Dahl (1992) for exploring the interaction between bank capital levels and asset risk, Van Roy (2005) analyzes adjustments in capital and credit risk at 576 banks in six G-10 nations between 1988 and 1995. In an effort to control for country and bank fixed effects, Van Roy includes country dummies and bank and country disturbances in simultaneous regressions in which capital and a measure of asset risk are interdependent dependent variables. Among various control variables, he includes a measure of "regulatory pressure" intended to reflect the degree of bindingness of the Basel I capital requirements. He finds evidence that low-capital banks in Canada, Japan, the United Kingdom, and the United States responded to this measure of regulatory pressure by increasing their capital but that low-capital banks in France and Italy did not.

Shrieves and Dahl (2003) study earnings management by Japanese banks within the 1989–1996 period of financial distress in that nation's banking system. They find evidence that Japan's banks were capital constrained during this interval and that the banks managed reporting of loan-loss provisions and gains on sales of securities in ways that smoothed reported income and replenished regulatory capital. Thus, Japanese banks engaged in regulatory-capital arbitrage that enabled them to satisfy capital requirements largely through earnings management rather than via direct balance-sheet adjustments.

Barrios and Blanco (2003) develop partial-adjustment models of bank capital in response to market forces versus capital constraints. They estimate these alternative

partial-adjustment frameworks using unbalanced annual panel data for seventy-six Spanish commercial banks between 1985 and 1991. Barrios and Blanco find that for their sample of banks, the market-based model better fits the data, indicating that the banks they considered were not at all constrained by capital regulation during the period of study.

Utilizing a buffer-adjustment approach to bank capital and in an analysis of data for 570 German savings banks between 1993 and 2004, Stolz (2007) finds that banks with relatively lower capital levels exhibit a negative relationship between the amount of capital and the degree of asset risk. In contrast, she finds a positive relationship between capital levels and asset risks at banks with greater capital buffers.

Beatty and Gron (2001) examine data for 438 publicly traded U.S. bank holding companies between 1986 and 1995. For the entire sample as a whole, their analysis suggests that pre-and post-regulation behavior of the entire set of banks was not materially affected by the advent of risk-based capital regulation.

Jackson et al. (1999) review a number of earlier studies investigating how capital adequacy regulations influence actual capital ratios, such as Peltzman (1970), Mingo (1975), Dietrich and James (1983), Shrieves and Dahl (1992), Keeley (1988), Jacques and Nigro (1997), Aggarwal and Jacques (1997), Hancock and Wilcox (1994), Rime (2001), and Wall and Peterson (1987). Jackson et al.'s conclusion is that there is little conclusive evidence that capital regulation has induced banks to maintain higher capital-to-asset ratios than they otherwise would choose if unregulated. Jackson et al. do conclude, however, that on balance most evidence suggests that in the near term banks mainly respond to toughened capital requirements by reducing lending.

In a creative study, Furfine (2001) develops an intertemporal model that yields optimizing conditions that guide bank decision-making with and without capital regulation. His approach is then to utilize data from FDIC call reports for 362 banks with assets exceeding $1 billion to estimate these optimizing conditions directly. Furfine then utilizes these estimated conditions to conduct simulations of the effects of changes in capital requirements. Based on his simulation analysis, Furfine's conclusion is that while capital regulation does matter, toughened supervisory scrutiny that accompanies explicit capital requirements generally has a larger influence banks' balance-sheet choices. Thus, he finds that the joint effects of tighter capital regulation and heightened supervision were more likely to have explained responses to the initial implementation of the Basel Accord in the early 1990s.

Most work points to this particular period as perhaps the most likely example of a regulatory-induced shock to aggregate credit. Nevertheless, Berger and Udell (1994) examine whether the risk-based capital requirements put into place in the late 1980s contributed to the so-called "credit crunch" that occurred in the United States in the early 1990s. They find evidence that other sources of loan supply reduction or declines in loan demand during the early 1990s played much more prominent roles in reducing bank lending. In contrast, Peek and Rosengren (1995a, b) conclude that there is considerable evidence, at least for New England, that both lower loan demand and a capital-crunch-induced decline in loan supply *together* brought about

a decline in lending. Brinkmann and Horvitz (1995) also find evidence of significant loan supply responses to the Basel I capital requirements.

Chiuri et al. (2002) extend the approach of Peek and Rosengren (1995a, b) to examine a panel of data for 572 banks in 15 developing countries. They find consistent evidence—after seeking to control for banking crises that occurred in 10 of the nations—that the imposition of capital regulation induced a reduction in loan supply and, hence, total lending in these countries. Furlong (1992), Haubrich and Wachtel (1993), and Lown and Peristiani (1996) also conclude that capital regulation contributed to a decrease in lending that helped fuel a post-capital-requirements U.S. credit crunch. Wagster (1999) reaches the same conclusion for Canada and the United Kingdom. He fails to find support, however, for this result in the cases of Germany, Japan, and the United States, where he joins Berger and Udell in concluding that a number of factors played a role in generating a credit crunch.

On net, therefore, the evidence regarding shock-producing effects of tougher capital regulation is mixed. There is neither full agreement that capital regulation has generated higher bank capital ratios nor that higher levels of capital necessarily translate into reduced risks. To the extent that increases in capital requirements have contributed to higher capital ratios, other factors appear to have figured into shocks created by a tightened supply of credit.

Procycical Features of a Capital-Regulated Banking Industry

The public demand for credit and supply of deposit funds to banks are positively correlated with variations in economic activity. Hence, banking inherently tends to be a procylical industry. As noted by Goodhart, Hofmann, and Segoviano (2004), financial liberalizations during the past two decades in many of the world's nations have added to banking procyclicality. Relaxations of various controls on loan and deposit interest rates, credit allocations, and cross-border flows of funds have allowed bank credit supply and deposit demand to respond positively to variations in economic activity along with the public's credit demand and deposit supply. Thus, today economic booms (busts) naturally tend to engender greater increases (decreases) in equilibrium quantities of deposits and lending than in prior periods in which bank credit supply and deposit demand were restrained by various governmental controls.

As also noted by Goodhart et al., regulation also adds to the natural procylical tendencies of banking. The bank supervisory process tends to press banks to constrict lending during contractions in an effort to protect bank balance sheets from the risks generated by downturns. During upswings, in contrast, the supervisors tend to take a hands-off approach that leaves banks freer to expand credit.

A number of observers have suggested that the Basel systems of risk-based capital requirements enhance the procyclical effects of bank regulation. Indeed, under Basel II minimum capital will change alongside a bank portfolio's perceived riskiness. Hence, capital requirements could increase considerably during economic downturns that are associated with greater bank portfolio risks. One consequence of automatically stiffer capital requirements could be an enlargement of short-term

decreases in lending accompanying recessions. Another consequence could be higher market loan rates that would reinforce economic downturns. In this way, risk-based capital regulation can potentially add to banking procyclicality. Borio (2003) stresses factors that might limit the procyclical features of bank capital regulation, such as greater market transparency and more supervisory discretion—the other two pillars of Basel II—which he suggests could help mitigate the extent to which capital actually responds to downturn-induced boosts in risks. Nevertheless, the fact remains that the design of risk-based capital regulation adds to the natural and supervisory-enhanced procyclicality of banking operations.

Catarieneu-Rabell et al. (2005) suggest that the rating systems banks utilize could greatly influence procyclicality under the proposed Basel II system. Utilizing rating schemes that are more stable over cycles, they argue, such as those produced by external rating agencies, would not contribute to the procyclical tendencies of capital regulation. Catarieneu-Rabell et al. conclude that banks would have greater short-run profit incentives to adopt a rating system conditioned on the current point in the cycle, which would have an unintended effect of boosting procyclicality.

What is the potential for significant procyclicality to emerge as a byproduct of bank capital regulation? To examine this question, Estrella (2004b) examines a dynamic, forward-looking model in which a bank chooses an optimal path for its balance sheet based on its rational expectation of loan losses. He presumes that the mean path for these losses is already cyclical and shows via simulations that a risk-based capital requirement is likely to have procyclical effects, which he suggests can be partially offset only by "judicious calibration" of minimum capital requirements during a downturn.

In an effort to determine how bank capital charges would vary in responses to changing risks over the course of a business cycle, Kashyap and Stein (2004) estimate probabilities of U.S. loan default during the period from late 1998–2002. Simulations based on these estimates suggest a potential for capital regulation to generate greater procyclicality both for the banking system and, in particular, individual banks. Kashyap and Stein's simulations indicate that banks lending to lower-quality borrowers are less susceptible to cyclical influences, because they are already most affected by risk-adjusted capital regulation. In contrast, banks that regularly lend to higher-quality borrowers are more likely to experience procyclical capital adjustments as changes in economic conditions move initially less-risky assets into riskier classifications.

Goodhart et al. (2004) also undertake simulations in an effort to assess the potential for procyclicality under the forthcoming Basel II standards. These authors attempt to construct typical bank portfolios Mexico, Norway, and the United States, and they simulate how loan ratings of assumed unchanging loan portfolios would vary during economic booms and busts under both the standardized Basel II risk-weighting system and the IRB system that the Basel II agreement will require large institutions to utilize. They conclude that implementation of Basel II has the potential to considerably add to banking procyclicality.

Gordy and Howells (2006) conduct simulations of bank portfolio volatility under different rating systems and reach more sanguine conclusions. They argue that

"empirically realistic" rules guiding banks' reinvestment strategies, which would induce banks to identify and lend to higher-quality borrowers during economic downturns, considerably reduce the additional procyclicality associated with capital regulation. As noted by Goodhart et al., however, at low points of economic cycles banks realistically may be more hard-pressed than assumed by Gordy and Howells to locate a significant number of creditworthy borrowers.

Empirical Evidence on Procyclical Effects of Capital Regulation

In practice, has capital regulation actually proven to be procyclical? There is surprisingly little evidence on this question. Ayuso, Perez, and Saurina (2004) seek to address this question using panel data from the Spanish economy and banking system for the period from 1986 to 2000. Using a variety of business-cycle measures, they find evidence of an economically and statistically significant negative relationship between bank capital and the phase of the business cycle. Thus, Ayuso et al. conclude that, at least in Spain, bank capital regulation has tended to create procyclical movements in actual capital. In contrast, consideration of data from a panel of more than 2,600 German savings and cooperative banks leads Stolz (2007) to conclude that capital levels adjust countercyclically. In addition, Angkinand (2009) finds that during crisis period, nations that subject their banks to stricter capital requirements experience smaller output losses than nations that do not.

Bouvatier and Lepetit (2008) analyze a panel of 186 European banks over the interval from 1992 to 2004. They focus on loan loss provisions, which they argue are linked via credit risk management that is in turn affected by capital regulation. Loan-loss provisioning behavior, Bouvatier and Lepetit conclude, amplifies credit fluctuations.

Can any procyclicality inherent in risk-based capital regulation be reduced? Pennacchi (2005) suggests that the procyclical features of risk-based capital regulation can be offset via risk-based deposit insurance. Pennacchi proposes that regulators could require less adjustment in bank capital to a recession-induced increase in asset risk, with banks required instead to pay higher deposit insurance premiums via a system involving a moving average of deposit insurance contracts. Of course, increased deposit insurance premiums would induce declines in deposits that can also generate a decrease in lending, although this effect would be offset somewhat by somewhat lower insurance premium payments on the lower deposit base. Nevertheless, Pannacchi concludes that the result would be less procyclicality than under capital regulation alone. He utilizes U.S. bank data for the period 1987–1996 to illustrate how adoption of such a scheme could smooth out cyclical influences, thereby dampening the inherent procyclicality of risk-based capital regulation.

Pederzoli and Torricelli (2005) offer an alternative suggestion for reducing the procyclical tendencies of bank capital regulation. They propose adjusting capital requirements based on *ex ante anticipations* of asset risk changes instead of ex post *observations* of altered risks. Even though procyclical effects would remain

under their proposal, Pederzoli and Torricelli argue using 1971–2002 U.S. data that forward-looking adjustments of risk-based capital requirements would considerably smooth peaks and troughs of the business cycle.

Summary: Capital Regulation, Bank Behavior, and Market Structure

- There is widespread agreement in the banking literature that the immediate effects of constraining capital standards are a substitution away from lending to holding alternative assets and accompanying increases in market loan rates. There is also broad concurrence that the longer-term effect of capital regulation is likely to be an increase in capital ratios, which may or may not be accompanied by a rise in total lending. Taken together, the agreement about these two sets of conclusions indicates that risk-based capital requirements can bring about one oft-expressed objective: increasing the relative size of the 'capital cushion' protecting depositors and deposit insurers from losses in the event of isolated or widespread bank failures.
- The theoretical literature on bank capital regulation offers widely divergent conclusions about how risk-based capital requirements influence choices that banks *make on the margin*. This issue is central to determining whether risk-based capital regulation actually makes individual banks and the banking system as whole "safer," because a relatively enlarged capital cushion may nonetheless be rapidly dissipated if banks respond to capital regulation by making riskier asset choices or failing to devote sufficient resources to containing adverse selection or moral hazard risks.
- Predicted effects of capital regulation on a bank's marginal decision-making are sensitive to the analytical banking framework researcher employs. Taken together, the conclusions of studies emphasizing the role of banks as portfolio managers offers qualified support for the idea that risk-based capital regulation could induce some banks to hold less risky asset portfolios but other banks to select riskier portfolios, thereby leaving ambiguous the predicted effects on the overall safety and soundness of the banking system.
- Banking theories emphasizing deposits' role as demandable debt and the potential for deposit markets to be imperfectly competitive generally offer greatest support for safety-and-soundness enhancements solely through imposition of capital requirements. This conclusion results mainly from the emphases that these theories place on the potential for bank runs and adverse effects of excess competition in deposit markets and from a de-emphasis on the risk implications for bank asset portfolios.
- Models that reverse the relative emphases placed on the asset and liability sides of the balance sheet and that consider banks' role as monitors for moral hazard risks in lending usually are more likely to indicate that capital regulation does not necessarily improve bank safety and soundness. Some theories in this vein indicate that capital regulation alone cannot necessarily produce a regulator's

preferred outcome unless regulators supplement capital requirements with supervisory or market discipline, thus supporting the view espoused by the framers of the Basel II standards that capital requirements complement other forms of regulation. Nevertheless, some theories imply that capital requirements and other types of regulation may be substitutes instead of complements, implying perhaps that in some circumstances capital regulation could prove to be superfluous or even harmful to overall safety and soundness.

- In light of the various ambiguities in the theoretical literature on bank responses to capital regulation, it must be concluded that the intellectual underpinnings for the proposed Basel II system are not particularly strong. More than three decades of research have revealed a host of factors that can both buttress and undercut the theoretical case for the use of capital requirements to discourage excessive risk-taking by representative model banks. Yet only in recent years have researchers begun to examine the manner in which capital requirements affect the safety and soundness of a banking system populated by diverse institutions instead of representative banks. Until more headway is made on this crucial issue, regulators may wish to contemplate alternative approaches to bolstering the safety and soundness of the banking system rather than dramatically expanding the scope and complexity of the present capital-requirement superstructure.
- Aggregate bank capitalization relative to total assets and risk-weighted assets increased noticeably during the 1990s, suggesting on the surface a line of causation beginning with the implementation of the Basel system of capital regulation. Indeed, some studies find evidence that implementation of the Basel I requirements may have contributed to a credit crunch in the early 1990s. A number of studies, however, find less compelling evidence of a link between capital regulation and credit curtailments, and some researchers conclude that market forces may have played a greater role in bringing about the overall increase in bank equity. Studies are somewhat mixed in their conclusions about just how much of an effect capital requirements had on the aggregate bank equity upswing, but the general theme running through most work is that capital regulation probably contributed only slightly to the rise in overall bank equity.
- The bulk of research on the topic suggests that capital regulation likely enhances procyclical tendencies already present in the banking industry. Most efforts to evaluate the extent to which capital requirements add to banking procyclicality have focused on simulations of model banking systems, and the vast majority of these studies indicate considerable potential for capital regulation to have procyclical impacts. The little empirical work examining actual data indicates that to the extent it has been binding for the banking system as a whole, capital regulation likely has had at least mildly procylical effects. Researchers have offered suggested policy changes that might dampen the procyclical tendencies of the Basel I and Basel II systems of capital requirements, but so far no action to reverse procylical features of the Basel II framework has been taken.

Chapter 8
Market Discipline and the Banking Industry

During past decades, there has been an historical oscillation of financial stability policies across objectives of management, resolution, and prevention (Bordo, 2003). Following three decades in which national governments and regulators placed primary weight on the management and resolution of financial crises, the primary objective during the early 2000s was placed on prevention, culminating in the international bank regulation and supervision framework that has come to be known as Basel II. Although the collapses and near-collapses of some of the world's largest banking institutions have already led to calls for a revamped, "Basel III" regulatory framework, Basel II seems likely to be the basis for whatever international system of regulation that may eventually emerge during the 2010s.

Basel II's framers suggest an analogy in which the framework is an elevated foundation resting on three "pillars." Pillar 1 is a redesigned system of risk-based capital requirements. Pillar 2 is a guideline for supervisory review, and Pillar 3 is a set of rules intended to promote market discipline. More than 80 percent of the description of the new framework (Bank for International Settlements, 2006). Hence, researchers so far have shone the brightest spotlight on capital adequacy regulation.

There is a growing understanding, however, that the three-pillar foundation is important. Basel II (or, for that matter, a revamped Basel III) is unlikely to provide banking systems with either a level playing field or a stable foundation unless each pillar is sufficiently well designed and structured to hold. The presumption in most discussions of the Basel II pillars is that in practice each will prove to reinforce the others. This chapter evaluates this official presumption. It provides a basic overview of the Basel II guidelines regarding market discipline and related conceptual issues. It reviews alternative suggestions for contributing to improved bank safety and soundness via enhanced market discipline, including proposals mandating the issuance of subordinated debts. It discusses recent research that encompasses analysis of market discipline within a basic banking model and applies this model to an analysis of the relationship between bank market structure and market discipline. The chapter then considers what the literature tells us about the extent to which markets actually discipline banks and about interactions between market discipline

D. Van Hoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-642-02821-2_8, © Springer-Verlag Berlin Heidelberg 2010

and supervisory discipline applied by government regulators of the banking industry. The chapter concludes by evaluating the market discipline pillar in relation to the capital standards and supervisory process pillars within the Basel II framework.

The Market Discipline Pillar of Basel II

The basic structure of Basel II is laid out in the Bank for International Settlements' "International Convergence of Capital Measurement and Capital Standards" (2006). Roughly 150 pages lay out details of the Pillar 1 system of capital requirements. The BIS document devotes fewer than 40 pages to discussion of Pillars 2 and 3, however. The following is a summary of the essential aspects of the market discipline pillar.

According to the Bank for International Settlements (2006, pp. 226–228), the objective of the market discipline pillar of the Basel II framework is to "encourage market discipline by developing a set of disclosure requirements which will allow market participants to assess key pieces of information on the scope of application, capital, risk exposures, risk assessment procedures, and hence the capital adequacy of the institution."

While "in principle, banks' disclosures should be consistent with how senior management and the board of directors assess and manage the risks of the bank," there are no explicit provisions for penalizing banks that fail to disclose information. Instead, the BIS document admits "that supervisors have different powers available to them to achieve the disclosure requirements":

> There are a number of existing mechanisms by which supervisors may enforce requirements. These vary from country to country and range from "moral suasion" through dialogue with the bank's management (in order to change the latter's behavior), to reprimands to financial penalties.

Furthermore, "a bank should decide which disclosures are relevant for it based on the materiality concept," namely, that "information should be regarded as material if its omission or misstatement could change or influence the assessment or decision of a user relying on that information for the purpose of making strategic decisions."

Thirteen pages of tables outline requirements for the release of qualitative and quantitative information on a bank's capital structure and capital adequacy, to be produced to satisfy a bank's Pillar 1 capital requirements. The text concludes by suggesting that most banks should release information on a semi-annual basis but "large internationally active banks and other significant banks' should release it quarterly."

As to other disclosures, the BIS states that "banks should have a formal disclosure policy approved by the board of directors." This policy should address "the bank's approach for determining what disclosures it will make and the internal controls over the disclosure process." Furthermore, the BIS suggests that "banks should implement a process for assessing the appropriateness of their disclosures, including validation and frequency."

The Channels of Market Discipline

At the core of the market discipline pillar of Basel II is the disclosure of information by banks, particularly in relation to their capital positions. National banking supervisor are responsible for ensuring that banks establish policies for information disclosure either semi-annually or quarterly.

Left unaddressed is whether Basel II's market discipline provisions are consistent with the idea of market discipline as understood by academic researchers, banking practitioners, and bank depositors and investors. To what extent are the disclosure guidelines of Basel II likely to promote "discipline" by the market? And how much are the guidelines likely to promote a safer and sounder banking system? Let's look at the literature.

Lane (1993) describes market discipline as the provision of signals, typically in the form of cutbacks in sources of funds banks use to finance their asset portfolios, which give banks incentives to engage in solvency-promoting actions. Banks obtain funds by issuing deposits, other debt that typically is subordinate to deposits in the event of failure, and shares of equity ownership that are subordinate to both deposits and other debt. Market discipline thereby involves actions on the part of depositors, other debt holders, or equity owners to induce banks to undertake actions consistent with promoting solvency.

Motivations of Agents Who Discipline Banks

Flannery (2001) reviews key aspects of the market discipline process. Depositors, he notes can respond to information on bank distress by reducing funds held in deposit accounts; that is, they will maintain present deposit levels only if a higher market deposit return is forthcoming or cut back deposit holdings at an unchanged market deposit rate.

Bliss (2004) reviews how various uninsured agents are involved in the application of market discipline to resolving the principal-agent problems faced by financial firms. As he notes, the fact that holders of debts issued by a bank, such as subordinated notes, receive fixed returns or returns based on other short-term market rates means that these parties bear considerable credit risk. Consequently, holders of bank debts have a strong incentive to press for banks to release information about the nature of their risks. Because returns on debt instruments issued by a bank are fixed, market prices of these instruments vary with perceived changes in the bank's exposure to risks. When the bank's debts mature, current and potential new holders of the instruments can re-evaluate the risk position of the bank and either seek a readjustment of the return or the amount of additional debts they may choose to hold.

Because bank equity owners gain if the bank experiences gains from risk taking, owners of equity face more conflicted motivations regarding the bank's exposure to risk. These agents consider expected profits and risk simultaneously and take into account the responses of debtholders to the bank's riskiness. Individual investors can respond to what they regard as excessive risk taking by holding current equity shares only at lower prices or by reducing equity ownership at current prices. As

stockholders, they can also pressure management to alter the structure of the bank's balance sheet to rein in risks. Equity owners may even replace managers whom they have determined seek to act in their own interests instead of in the interests of the equity owners.

Conditions for Market Signals to Effectively Discipline Banks

The above discussion suggests that suppliers of funds to banks generally can react to a changes in the risk composition of a bank's balance sheet in one of two ways. One possible response could occur along the *intensive margin*. That is, if depositors, debt holders, or equity owners perceive greater bank solvency risk, they can reduce the amount of funds supplied to the bank or continue to supply the same quantity but only at a higher rate of return. A second possible response is along the *extensive margin*. That is, they can entirely cut off their supply of funds. Either of these forms of market influence would be what Kwast et al. (1999) refer to as "direct" influence.

Flannery (2001) notes that market influence cannot take place unless another part of the market discipline process is effective: Suppliers of funds to banks must be able to engage in *market monitoring*. They must be able to evaluate changes in banks' risk characteristics and financial conditions. Kwast et al. (1999) suggest that to the extent that information derived from market-monitoring activities by funds suppliers and perhaps even regulators feeds back to affect prices of secondary debt instruments issued by banks, "indirect" market influence may induce corrective action on the part of banks.

Several conditions must be satisfied for depositors, holders of bank debt and owners of bank equity both to perceive bank weakness and through their market responses to generate corrective action. First, they must have correct information at appropriate times. To assess the implications of that information, they must have information about bank management, capital structure, and risk exposures.

Second, if they are to provide appropriate market signals, suppliers of funds must have an incentive to react to perceived changes in the probability of bank insolvency; that is, suppliers' own funds must be at risk in the event of bank failure. Depositors, debt holders, and equity owners must not believe that banks will be bailed out by regulators, which requires in turn that regulators must be able to credibly commit not to bail out uninsured investors.

Third, the market signals to a bank encountering a perceived change in its creditworthiness must be visible to all interested parties, including all depositors, debt holders, and equity owners. This means that the markets in which suppliers of funds to banks participate—markets for bank deposits, subordinated debt, and equity—must be open and active. When suppliers of funds perceive a deterioration in a particular bank's solvency position and respond along intensive or extensive margins, the amount of funds supplied will decline, generating both a reduction in funding and an increase in the market rate that this bank must pay to obtain funding.

Finally, an affected bank must respond to these market signals in a solvency-promoting manner. This means that there must be incentives for a bank's managers to respond by reducing risk exposures or for equity owners to replace the affected bank's managers with new managers better attuned to failure risks.

Potential Benefits and Costs of Market Discipline in Banking

In a particularly comprehensive review of market discipline in banking, Hamalainen, Hall, and Howcroft (2003) enumerate several potential social benefits of effective market discipline. The most obvious is fewer moral hazard problems arising from government insurance guarantees to depositors. Another is improvements in cost efficiency through persistent pressures for the most effective management.

If private recognition and response times are faster than those of regulators, market discipline might aid regulators in differentiating "bad banks" from "good banks" and prompt them to move more quickly to seek resolution in the event of imminent failure. In principle, if market discipline acts as a check on management, the burden of proof that a specific bank is not experiencing difficulties would be shifted from regulators to managers, and society could reduce the regulatory burden imposed upon the banking system as a whole.

There are also potential negative by-products associated with market discipline of the banking industry. If most suppliers of funds to banks see significantly higher failure probabilities, they might respond along extensive margins, resulting in socially costly bank runs ex post (although the potential for runs can be ex ante a positive source of market discipline). If sophisticated savers have an advantage in observing and reacting to market information, unsophisticated savers might be at greater risk of incurring proportionately greater costs in the case of failures. Larger banks with more resources may be better able to provide the information required to permit market discipline. Hence, they could have a cost advantage over smaller banks—although it is unclear if such a size advantage exists or if so whether it is more pronounced in banking than any other industry. Banks in more developed nations with open and active public debt and equity markets could have advantages over banks in less developed countries with relatively closed and inactive markets for debt and equity. Finally, to the extent that suppliers of funds have persistently faulty perceptions and thereby transmit false market signals, other market participants and regulators may have misguided reactions.

Evaluating Incentives for Information Disclosure

Clearly, effective market discipline in banking requires market participants to be informed about banks' risk profiles. Immediate knowledge of risks, however, is private information to banks' managers, and theory suggests that full disclosure is unlikely to be a privately optimal outcome. In Verrecchia's (1983) model, for instance, investors recognize that the release of information by a firm's managers could be costly for the firm and thus are willing to accept some discretion on the part of managers to withhold a certain "threshold" quantity of information.

In addition, Boot and Thakor (2001) provide an analysis suggesting that the quantity and type of information a firm chooses to disclose depends on whether the private managerial information complements or substitutes for public information already possessed by investors. Complementary information disclosure,

they conclude, strengthens private investors' incentives to acquire information, while substitute disclosure weakens these incentives. Separately, Boot and Schmeits (2000) point out that the effectiveness of market discipline both influences and is itself influenced by the degree of conglomeration within financial institutions. They suggest that effective market discipline reduces the gains from conglomeration, while increased conglomeration tends to undermine the benefits of market discipline.

Furthermore, Östberg (2006) notes that empirical evidence indicates that disclosure of information typically increases a firm's market value, yet he suggests that it may not be in the interest of those controlling the firm to fully disclose all information if they want to expropriate a portion of investment returns. In Östberg's framework, therefore, a supervisor contemplating a policy mandating disclosure faces a trade-off: Weaker disclosure standards encourage socially costly expropriations by firm insiders, but stricter disclosure standards reduce the returns of the insiders and thereby discourage their investment. Thus, mandatory disclosure rules potentially could have counterproductive effects in Östberg's model.

Indeed, Rochet and Vives (2004) point out that too much disclosure could trigger systemic banking crises. Chen and Hasan (2006) develop a theoretical model in which greater information transparency on the part of banks tends to boost the likelihood of bank runs, unless bank disclosures clarify to depositors that the bank's problems are idiosyncratic rather than systemic. Recent work by Huang and Ratnovski (2008) focuses on the significant role of "wholesale funds—that is, large-denomination time deposits and various purchased funds, such as repurchase agreements and Eurocurrency deposits—as sources of funds for banks." They build on the Calomiris and Kahn's (1991) analysis discussed in Chapter 6. Whereas Calomiris and Kahn show how sophisticated sellers of wholesale funds can add value by monitoring banks and imposing market discipline, Huang and Ratnovski show that under some circumstances, the presence of a noisy public signal on the quality of project quality can induce sellers of wholesale funds to acquire less information and engage in less efficient monitoring. As a result, receipt of an unexpectedly negative signal can generate significant withdrawals of wholesale funds, resulting in socially inefficient bank liquidations of projects.

Ways to Enhance Bank Market Discipline

The corporate governance literature suggests that in most industries investors are well placed to control or influence borrower behavior in primary markets. Distortions of market-based governance incentives by government guarantees or regulatory mechanisms can, however, weaken the ability of investors to discipline firms. Even in a world where investors have full information and market signals are clear to all participants, markets will fail to discipline banks if government policies insulate managers from the signals that markets transmit to them. In banking, deposit insurance and regulatory policies such as a too-big-to-fail doctrine likely

account for the conclusions reached by Bliss and Flannery (2000) as well as for documented variations in market reaction to information under different regulatory regimes.

Because depositors can obtain deposit insurance coverage by holding their deposits in different institutions, researchers have focused considerable attention on subordinated debt markets as a source of market discipline for both public and privately held banks. As Saunders (2001) notes, in principle markets for bank equity and subordinated debt should provide equally useful information. Indeed, Allen and Gottesman (2006) find evidence supporting the integration of bank equity and bond markets.

As Saunders observes, however, explicit U.S. legal changes in the early 1990s expressly exempted subordinated debts from too-big-to-fail guarantees. Thus, suppliers of funds to banks via these debt markets should be particularly responsive to information about bank risks, and bondholders in these markets have considerable downside risks. To the degree that banks rely on subordinated debt issues to fund their asset portfolios, the markets for these debt instruments could well be a potential channel for market discipline even in the world's heavily regulated banking systems. Consistent with this presumption, Ashcraft (2008b) finds evidence suggesting that greater shares of regulatory capital allocated to subordinated debt are associated with a greater likelihood that distressed banks will experience positive outcomes.

Maclachlan (2001, p. 228) suggests that "the idea of market discipline operating in highly regulated and protected industry seems somewhat paradoxical." Nevertheless, many authors have floated a number of proposals for making the issuance of bank subordinated debt mandatory to enable regulatory use of the information revealed by subordinated debt yield spreads (see, for instance, Evanoff and Wall, 2000, Calomiris, 1999, 2004, Kwast et al., 1999, and Bliss, 2001). Among the features included in typical mandatory subordinated debt proposals are the following:

1. The inclusion of no-bailout clauses in debt contract provisions to ensure debt holders have strong incentives to monitor bank risk profiles.
2. Restrictions on holdings by bank insiders.
3. A requirement for subordinated debt instrument maturities to be long enough that perceived failure risks are priced but short enough that bank must go to the market regularly to roll over debt; most proposals suggest maturities of one to five years.
4. A requirement of staggered issue dates to ensure a significant number of diverse debt holders at any given time and thereby make a consistent market signal more likely.
5. A requirement for debt instruments to be issued in minimum denominations large enough to ensure holders will have strong incentives to monitor risks.
6. Integration into the Basel capital requirements regime to raise the profile of subordinated debt for disciplinary purposes.
7. The inclusion of corrective action rules, such as debt covenants imposing stricter sanctions as bank performance deteriorates (which Goyal (2005) finds

are more common in subordinated debt issues when overall banking industry risks increase); prompt corrective regulatory response to a bank's inability to issue new debt (see, for instance, Llewellyn and Mayes, 2003); and puttable debt arrangements permitting debt holders to exercise put options on issuers at any time, triggering regulatory response.

One rationale for mandatory subordinated debt issuance is the suggestion by Lang and Robertson (2002) that such an approach would replace regulators' subjective judgment with market signals at not much of an increase in the risk sensitivity of bank costs. Calomiris and Powell (2000) argue that this approach would eliminate "plausible deniability" for regulatory inaction, thereby giving bank supervisors less latitude for forbearance in dealing with failing banks.

Blum (2002) provides a theoretical analysis of bank risk-taking behavior in a model with subordinated debts. Bank managers choose the level of portfolio risk subsequent to the determination of the rate of return paid on these debts. As a consequence, realized portfolio risk depends on whether the bank can precommit in advance to a given level of risk. In a commitment equilibrium, issuing subordinated debts reduces bank risks, but absent an ability to precommit portfolio risk increases.

In addition, Niu (2008) extends Repullo's (2004) framework by including subordinated debt. In this amended version of Repullo's model, holders of subordinated debts must try to anticipate a bank's portfolio choice and determine the rate at which they are willing to lend to the bank equity owners' choice of an asset portfolio that maximizes their expected returns. Niu's analysis suggests that for a level of subordinated debts below a critical threshold, increasing subordinated debt decreases the capital requirement necessary to reduce a bank's risk-taking incentives.

Nevertheless, mandatory subordinated debt proposals raise almost as many issues as they are intended to resolve. Hancock and Kwast (2001), for instance, find that subordinated debt spreads are most consistent across alternative sources of data for the most liquid U.S. bonds, which typically are those issued in relatively large amounts and by numerous medium to large-sized U.S. banks. Yet Sironi (2001) documents that trading in European markets for subordinated debt tends to be heavily concentrated among issues of the largest European banks and that the secondary market for these instruments is relatively illiquid. Undoubtedly, such markets are even thinner in less developed nations. Thus, it is unclear whether market discipline is likely to be as effective everywhere as it might be in more diverse and liquid U.S. markets.

Park and Peristiani (2007) also point out that bank stockholders can either be "enemies" of regulators, by condoning increased risk-taking on the part of banks with option values that outweigh their charter values, or "allies," by penalizing risky strategies of low option value institutions. As a bank's risk of failure increases, equity holders are more likely to switch from ally to enemy. Using Tobin's Q to infer bank charter values, Park and Peristiani estimate that a switching point for U.S. banks during the 1986–1992 period occurred when there was an annual failure probability of 17 percent.

As noted by Herring (2004), another potential problem associated with mandated subordinated debt proposals is that integrating market discipline into regulatory requirements on banks has the potential to increase regulatory compliance costs. Sophisticated systems might be required to evaluate the nature of the risks implied by market signals, implying higher fixed costs for banks. Furthermore, banks could confront significant costs of making regular marginal balance-sheet adjustments in response to variations in market signals.

The corrective action feature in mandatory subordinated debt proposals raises a crucial issue: Assuming that signals in the markets for bank liabilities do reveal useful information for disciplining banks, how should bank regulators use the information? Should corrective action be immediate or gradual? Immediate supervisory actions make immediate failure more likely, while progressively stricter sanctions might give managers the time to save a weakened bank. A more gradual approach to corrective action could simply give the managers more time to weaken the institution further and increase closing costs for society; if so, some kind of rule-based corrective action triggered by market signals might be appropriate.

Industry Structure and Market Discipline

Most work on market discipline in banking assumes that most bank assets and liabilities trade in competitive markets. Only recently have a few researchers explored the role of industry structure as a factor influencing market discipline.

Market Discipline in a Basic Banking Model

To consider issues relating to effects of market discipline on bank behavior, consider the following version of the model proposed by Landskroner and Paroush (2008). An individual risk-neutral bank maximizes, subject to its balance-sheet constraint $L^i + S^i = (1–q)D^i + N^i + E^i$, its expected end-of-period value of equity, $\hat{E}^i$, given by

$$\hat{E}^I = P(E^i)\Pi - C^i(L^i, S^i, D^i, N^i, E^i).$$

where E^i denotes the bank's beginning-of-period equity capital, which is assumed to be unregulated to facilitate an evaluation of the essential equity incentives a bank confronts from market discipline, and where $\Pi = R^i_L(L^i)L^i + R_S S^i - R^i_D(D^i)D^i - R^i_N(N^i,E^i)N^i$ defines gross interest earnings net of gross interest expenses, with $R^i_L = 1 + r^i_L, R_S = 1 + r_S, R^i_D = 1 + r^i_D$, and $R^i_N = 1 + r^i_N$.

In the above expression, $P(E)$ is the probability of receiving a non-zero stream of gross interest income, where $0 < P(E^i) < 1$, $P'(E^i) \geq 0$, and $P''(E^i) \leq 0$. An increase in the bank's beginning-of-period equity capital is assumed to have a non-negative impact on $P(E)$, thus improving the likelihood that the bank will obtain sufficient end-of-period resources derived from net interest income to avoid a negative market value after accounting for noninterest expenses. The model allows for

the bank potentially to possess some degree of market power in the markets for loans, deposits, and non-deposit liabilities, where $R_L^i(L^i)$, $R_D^i(D^i)$, and $R_N^i(N^i,E^i)$ are the inverse loan demand and deposit and non-deposit-liability supply functions, with $\partial R_L^i/\partial L^i \leq 0$, $\partial R_D^i/\partial D^i \geq 0$, and $\partial R_N^i/\partial N^i \geq 0$. In addition, $\partial R_N^i/\partial E^i < 0$; hence, if the bank raises its beginning-of-period equity, the bank generates a reduction in the rate of return it must pay to obtain non-deposit-liability funds.

The following condition holds true when the bank maximizes its expected end-of-period equity:

$$\begin{aligned} P(E^i)[1-(\varepsilon^i)^{-1}]R_L^i - C_L^i &= P(E^i)R_S - C_S^i = P(E^i)[1+(\mu^i)^{-1}]R_N^i + C_N^i \\ &= (1-q)^{-1}\left\{[1+(\eta^i)^{-1}]P(E^i)R_D^i + C_D^i\right\} \\ &= P'(E^i)\Pi + P(E^i)R_N^i\left(1-\xi^i\frac{N^i}{E^i}\right) + C_E^i, \end{aligned}$$

where ε^i is the absolute elasticity of demand for loans, η^i is the absolute elasticity of supply of deposit funds, μ^i is the absolute elasticity of supply of non-deposit-liability funds, and $\xi^i \equiv \frac{(\partial R_N^i/R_N^i)}{(\partial E^i/E^i)}$ is the absolute elasticity of the rate of return on non-deposit liabilities with respect to the bank's beginning-of-period equity capital. Landskroner and Paroush (2008) propose ξ^i as a measure of the degree of market discipline imposed upon the bank. An increase in ξ^i implies a greater proportionate change in the market rate of return on non-deposit liabilities in response to a proportionate change in the bank's equity.

Note that because $\partial R_N^i/\partial E^i < 0$, an increase in beginning-of-period equity capital generates a reduction in the gross return that the bank must pay on its nondeposit liabilities. Thus, at an initial balance sheet configuration satisfying the above condition, the immediate effect of an exogenous increase in the value of ξ^i is a reduction in the marginal cost of beginning-of-period equity capital. Hence, an exogenous increase in the degree of market discipline boosts the incentive for the bank to substitute away from deposit and non-deposit liabilities in favor of additional equity, consistent with the conclusion obtained by Landskroner and Paroush. *Ceteris paribus*, the bank utilizes the additional beginning-of-period equity to expand its holdings of securities and its lending, which it brings about via a reduction in the market loan rate. If $P'(E^i) > 0$, the resulting rise in equity also raises the probability of attaining positive gross interest income and maintaining a positive end-of-period value of equity.

Market Power, Information Disclosure, and Market Discipline

Most research regarding the impact of market structure on information disclosure suggests that firms in less competitive industries have incentives to be less forthcoming. As discussed by Ali et al. (2008), theoretical work suggests that it is optimal for firms to offer less informative disclosure in oligopolistic industries in which incumbents have interdependent investment strategies that they desire to withhold

from rivals. Ali et al. provide evidence supporting this predicted relationship derived from U.S. manufacturing firms between 1995 and 2005.

Bolton et al. (2004) examine incentives for truthful revelation of information by financial intermediaries under specialized and universal banking arrangements. Under either structure, Bolton et al. conclude that under monopoly conflicts of interest may prevent information disclosure. Greater competition among intermediaries increases the degree of disclosure, with full disclosure emerging when banks are competitive and reputation costs are sufficiently great.

So far, the banking literature has had relatively little say about the effects of variations in market power on bank market discipline. This issue can be examined in the context of the Landskroner–Paroush framework, however. Based on the previous discussion of this model, note first that the spread between the rates of return on non-deposit liabilities—such as subordinated debts—and on deposit liabilities in terms of the exogenously determined market security rate and relative to the return on deposit liabilities can be expressed as

$$\frac{R_N^i - R_D^i}{R_D^i} = \frac{[P(E^i)R_S - C_S^i - C_N^i][1 + (\eta^i)^{-1}]}{\{(1-q)[P(E^i)R_S - C_S^i] - C_D^i\}[1 + (\mu^i)^{-1}]} - 1.$$

This rate spread is decreasing with a rise in the value of η_i or a reduction in the value of μ_i. Thus, consistent with an analogous result obtained by Landskroner and Paroush, a reduction in market power in the market for deposits narrows the spread between the rate on non-deposit liabilities and deposit liabilities. In this version of their model, which allows for market power in the market for non-deposit liabilities, an increase in the bank's pricing power in this market brings about a widening of this spread.

At an initial value-maximizing balance-sheet configuration and hence the corresponding initial value of the gross market return on non-deposit liabilities, the market loan rate is determined by

$$R_L^i = \frac{P'\left(E^i\right)\Pi + P\left(E^i\right)R_N^i\left(1 - \xi^i\frac{N^i}{E^i}\right) + C_E^i + C_L^i}{P\left(E^i\right)\left[1 + \left(\varepsilon^i\right)^{-1}\right]}.$$

This expression implies that an exogenous increase in the degree of market discipline, ξ^i, which induces the bank to expand its beginning-of-period equity, also gives the bank an incentive, as noted above, to reduce its loan rate as it seeks to allocate the additional equity to earning assets. The extent to which the bank reduces its loan rate in response to greater market discipline depends on the loan demand elasticity, ε^i. A larger value of ε^i reduces the loan rate markup, which in turn decreases the responsiveness of the loan rate to increased market discipline. Thus, a decline in bank market power cuts the influence of market discipline on the loan rate and, hence, on lending.

In a complete solution of this model, the degree of market discipline itself should emerge as an endogenous variable. Consider, for instance, the special case in which

$P'(E^i) = 0$, so that the probability of bank insolvency is independent from E^i. In this situation, the value-maximizing condition implies that at an optimal balance-sheet configuration:

$$P(E^i)[1 + (\mu^i)^{-1}]R_N^i + C_N^i = P(E^i)R_N^i\left(1 - \xi^i\frac{N^i}{E^i}\right) + C_E^i.$$

This in turn implies that at an optimum, the degree of market discipline is $\xi^{i*} = 1 - \left(\mu^i\right)^{-1}\frac{E^i}{N^i}$. Thus, given the level of market power the bank possesses in the market for its non-deposit liabilities—as governed by the supply elasticity μ^i—equilibrium market discipline is decreasing in the ratio of beginning-of-period equity to non-deposit liabilities. As the bank increases its issuance of non-deposit liabilities in relation to beginning-of-period equity, it expands its exposure to the disciplining impact of variations in the rate of return in that market, thereby pushing up the value of ξ^i—that is, raising the proportionate responsiveness of the return on non-deposit liabilities to a given change in equity—which in turn feeds back to induce an increase in equity. Thus at the degree of market discipline ξ^{i*}, the bank balances its issuances of beginning-of-period equity and non-deposit liabilities at a level that maximizes the end-of-period equity value.

Alternatively, for a given ratio of beginning-of-period equity to non-deposit liabilities, an increase in the magnitude of the supply elasticity μ^i raises ξ^{i*}. At higher values of μ^i, the bank's ability to mark down the rate on non-deposit liabilities declines, making this rate more responsive to variations in other variables, including the bank's beginning-of-period equity. There is an increase, therefore, in the value of the elasticity of the rate of return on non-deposit liabilities with respect to beginning-of-period equity. Less market power in the market for non-deposit liabilities thereby translates into a greater degree of market discipline, a result that accords with arguments advanced by Landskroner and Paroush. Indeed, in this particular version of their model, in the competitive limit $\mu^i \to \infty, \xi^{i*}$ equals unity. When the bank possesses no market power in its issuance of non-deposit liabilities, a given proportionate increase in its beginning-of-period equity elicits an equiproportionate reduction in the rate of return on those liabilities.

As Landskroner and Paroush demonstrate, the analysis becomes more complex and hence forthcoming predictions more ambiguous in the general case in which $P'(E^i) > 0$. Nevertheless, the basic result that greater competition in the market for non-deposit liabilities enhances market discipline generated by variations in the rate of return on such liabilities is likely to survive most model generalizations. The converse is also likely to hold true: Possession of greater market power in the market for non-deposit liabilities can be predicted to reduce the usefulness of variations in the rate of return on these liabilities as an instrument of market discipline.

Evidence on Market Discipline's Effectiveness

The potential for market discipline to complement or perhaps even serve as a substitute for bank supervision and regulation has long been recognized (see, for instance,

Gilbert, 1990). Given Basel II's emphasis on market discipline and capital regulation, however, we should look at the evidence regarding bank market discipline since the establishment of capital requirements under the first Basel Accord, or Basel I.

Information Content of Market Prices and Bond Yield Spreads under Basel I

A key issue is whether suppliers of funds really can perceive changes in banks' risk profiles. Flannery (1998) provides a detailed review of the evidence through the mid-1990s and concludes that there is a role for market discipline in supplementing regulatory supervision. Flannery and Sorescu (1996) indicate that investors can distinguish among risks taken by major U.S. banks, especially during periods when subordinate debt is not covered by government guarantees. They observe that spreads between yields on subordinated debt and yields on Treasury securities with the same maturities may have better leading indicator properties than capital adequacy measures in predicting bank condition.

Distinguin et al. (2006) use a logit early warning model to test whether accounting data and equity markets assist in predicting distress for European banks. They find some evidence that market-based indicators have predictive power for assessing the degree of bank stress, but only for banks that rely less heavily on insured deposits.

Flannery et al. (2004) find that stocks of large bank holding companies have similar trading properties to those of matched nonfinancial companies. Stocks of smaller bank holding companies, trade much less frequently but have spreads similar to those of matched nonfinancial companies. They conclude that bank holding company stocks appear to be no more informationally opaque than other stocks, although it is possible that regulatory supervision could promote transparency.

Morgan (2002) explores "bank opacity" by investigating bond rating disagreements between Moody's and Standard & Poor's. The data suggest that, as compared with other industries, the rating agencies disagree more sharply when evaluating the bond issues of banks. Bank loans and trading assets are a significant source of disagreement between the two bond raters. These factors, Morgan suggests, indicate that banking is characterized by more opaqueness than other industries. Building on Morgan's work, Iannotta (2006) studies ratings differences for bonds issued by European banks and reaches similar conclusions.

Morgan and Stiroh (2001) examine market spreads on nearly 500 new bonds issued by U.S. banks and bank holding companies in primary markets. They find that spreads widen with riskier asset mixes. This result, they suggest, indicates subordinated debt markets are able to price risk structures of bank asset portfolios, although banks are most likely to enter primary debt markets when investors' perceptions of bank risks are more favorable. Because these results apply to a period of especially good times for U.S. banks (1993–1998), Morgan and Stiroh conclude that debt markets provide clear signals of asset risk differentials across banks even when overall industry risk is relatively low.

In a different approach to evaluating spreads as indicators of bank risk, Evanoff and Wall (2001) compare the performance of capital ratios and subordinated debt spreads as predictors of regulatory ratings of banks. They find that debt spreads have significantly more predictive power than all capital ratios except for the Tier 1 ratio, where debt spreads are only marginally better predictors of future ratings by regulators.

Does a low predictive power of subordinated debt spreads necessarily imply that market discipline is ineffective? Bond et al. (2008) propose a negative answer in the context of a theoretical model in which agents making corrective-action decisions in light of new information about a bank utilize the market rate of return on subordinated debt as a source of information and, simultaneously, engage in actions that collectively influence that rate of return. They find that a possible consequence of the interaction between using the rate of return on subordinated debt as an information variable while concurrently acting on information about the bank's prospects can result in little net change in the equilibrium value of the rate of return, implying that an observation of little change in a subordinated debt rate spread cannot necessarily be interpreted as indicating a lack of market discipline.

Market Discipline versus Regulation

Several researchers have sought to determine whether market rates of return on bank liabilities convey more or less information than the traditional regulatory process. On the basis of an examination of Italian bank data, Cannata and Quagliariello (2005) conclude that bank stock price movements generally reflect information obtained in supervisory examinations. In an analysis of U.S. data, Berger, Davies, and Flannery (2000) find that, with the exception of on-site inspections, supervisory assessments are less predictive of future changes in performance than equity or bond market indicators. Indeed, a study by Iannotta (2006) of bond spreads for more than 100 European banks between 1999 and 2007 suggests that investors are able to discern some hidden information about the banks' performances and prospects.

Nevertheless, work by Berger and Davies (1998) and DeYoung et al. (2001) find that bank examinations provide new information that market prices do not immediately reflect, indicating that regulation may uncover information more readily than markets for bank debts. DeYoung et al. suggest that regulators are more likely to uncover "bad" private information. They also find that new information uncovered by supervisory examinations influences bank values (although Cole and Gunther (1998) offer evidence that the value of private information obtained in bank examinations vanishes within six months) and that subordinated debenture prices do not immediately reflect this information. DeYoung et al. argue that when information obtained by supervisory examinations is released, market prices reflect not only the information itself but also the likely regulatory actions implied by the information.

Should regulators release all information obtained from supervisory examinations? Prescott (2008) examines a model involving interactions among a bank, a

regulatory supervisor, and outside investors, in which a bank deciding whether or not to be honest with the supervisor balances the expected gain from lying to supervisors against the penalty from doing so. In the model environments he considers, which consider endogeneities in neither the bank's rate of return nor the size of the bank's investment, disclosure to investors of supervisory information either leads the bank to transmit lower-quality information or requires the supervisor to devote more resources to extracting information. As Prescott notes, another factor not taken into account in his framework is that public release of supervisory information could pressure a regulator to move more quickly to initiate prompt corrective action.

Regulatory Crowding Out of Market Discipline?

Billett et al. (1995) provide an analysis suggesting, however, that regulatory actions in response to new information about bank risks give imperfectly competitive banks incentives to reshuffle liabilities to protect themselves from market discipline. Figure 8.1 illustrates the argument of Billet et al. under the assumption of no reserve requirements. Initially, the bank maximizes profits by equating $[1 - \varepsilon^{-1}]\, r_L - C_L$, the net marginal revenue on loans in an imperfectly competitive loan market, with $r_N + C_N$, the marginal factor cost of nondeposit-liability funds assumed to be obtained competitively. The bank also equates $[1 + \eta^{-1}]r_D + C_D$, the marginal factor cost of deposit funds obtained in an imperfectly competitive deposit market, with $r_N + C_N$. These actions determine the volumes of deposits and nondeposit liabilities and the total of the two, which are assumed to constitute the bank's total liabilities.

The bank's compliance with regulatory toughening in reaction to the information of the bank's increased riskiness necessitates incurring higher marginal resource costs associated with lending (C_{L}), deposits (C_{D}) and non-deposit liabilities (C_{N}). In addition, the market rate of return on non-deposit liabilities increases as those holding these liabilities respond by requiring a higher risk premium. On net, *ceteris paribus*—in particular, assuming unchanged loan and deposit rates—the bank responds to the higher relative cost of nondeposit liabilities by substituting deposits for nondeposit liabilities while reducing its total liabilities. In most nations, a large portion of deposits are government-insured while nondeposit liabilities typically are not; thus, the bank responds by relying more heavily on primarily insured liabilities while cutting back on uninsured liabilities. The disciplining presence of the latter thereby shrinks. They conclude, therefore, that the regulatory response to rating downgrades together with the higher risk premium on nondepository liabilities leads to a decline in the scope for the latter to provide a source of market discipline.

Billett et al. examine data on 116 downgrades in credit ratings of US. bank holding companies between 1990 and 1995. They find that, consistent with the theoretical implications of Fig. 8.1, banks did indeed increase their absolute and relative issuances of insured deposits following rating downgrades, suggesting regulatory crowding out of market discipline acting through the presence of nondeposit liabilities on banks' balance sheets.

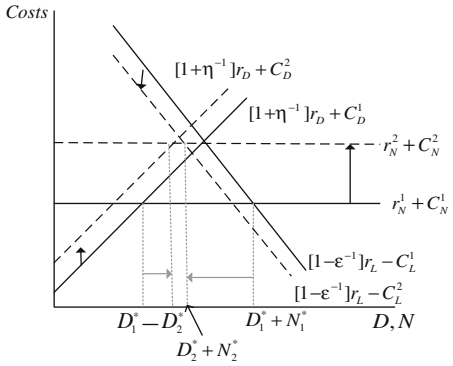

Fig. 8.1 Combined effects of increased bank risk with regulatory and market discipline reactions

King (2008) obtains a similar result in an analysis of the disciplining role of the market for interbank funds. Based on quarterly data on U.S. banks over the 1986–2005 period, King concludes that high-risk banks consistently pay higher interest rates on interbank loans and respond by substituting away from these nondeposit-liability sources of funds.

Likewise, in a study of deposit rates of 26 Russian banks between 2002 and 2005, Peresetsky et al. (2007) find that deposit rates rose in response to increases in measures of bank risk. Although their conclusions must be treated with some caution in light of a relatively small number of observations, Peresetsky et al. conclude as well that the September 2004 introduction of deposit insurance induced banks both to substitute heavily into insured deposits and to take on greater risk, suggesting a reduction in market discipline as a consequence of this regulatory change.

Additional Evidence on Interactions between Regulation and Market Discipline

Given the heavy regulatory presence in banking markets, it is a challenge to disentangle the effects of regulatory stances and practices on market discipline. Covitz, Hancock, and Kwast (2004) seek to do so by considering bank subordinated debt spreads relative to rates on comparable-maturity Treasury securities during three separate regulatory regimes: the too-big-to-fail period of 1985–1987; the purchase-and-assumption period of 1988–1992; and the post-FDIC Improvement Act period of 1993–2002. They find that spreads were responsive to bank-specific risks in all three regimes, but they became more sensitive to risks in the second. Spreads became less sensitive to risks in the third regime, perhaps indicating some perception that FDICIA provisions reduced downside risk for investors.

There is only limited evidence on whether uninsured depositors are able to perceive and respond to market signals of bank risks. Birchler and Maechler (2001)

examine Swiss data and find that variations in bank-specific fundamentals explain as much as 75 percent of the variation in a bank's uninsured deposits. Separately, Park and Peristiani (1998) estimate a logit model of U.S. thrift institution failures and find that riskier thrifts pay higher deposit rates and attract smaller amounts of uninsured deposit funds. In addition, McDill and Maechler (2003) model the behavior of uninsured deposits in the U.S. banking sector as an autoregressive process. They find that uninsured deposits in a variety of institutions respond to changes in fundamentals including the capitalization, relative size, and type of institution. Finally, in a case study of the Comptroller of the Currency's corrective actions against Hamilton Bank between 2000 and its ultimate failure in early 2002, Davenport and McDill (2006) document that business holders of uninsured deposits were particularly sensitive to news of the bank's deteriorating conditions—that is, the business supply of deposits was relatively more elastic.

Hagendorff et al. (2008) focus on the market for bank equity as a potential mechanism of market discipline. They examine, in the context of data for the 1996–2004 interval, how characteristics of bank boards of directors in the United States and Europe were associated with returns of acquirers following merger and acquisition announcements and with merged institutions' long-term financial performances. Hagendorff et al. find that more frequent meetings of bank boards in the United States are associated with higher post-announcement returns and longer-term improvements in performance than is the case in Europe. In light of their assertion that the U.S. bank regulatory regime is relatively tougher than in the United States, Hagendorff et al. interpret their results as implying complementarity of regulation and equity-market discipline in the United States but not in Europe.

Requiring banks to formulate processes for disclosure of information is at the heart of the Basel II market discipline pillar, which extends on the Basel I framework's emphasis on regular and accurate financial reporting. There is evidence that improved systems for the disclosure of financial data do appear to promote a safer banking system. For instance, Podpiera (2006) finds evidence in panel data from 65 nations during the 1998–2002 period that adherence to this and other "core principles" of the Basel I framework improved bank performance. In a study examining Moody's financial-strength rating for more than 200 banks in 39 countries, Demirgüç-Kunt et al. (2008) conclude that nations adhering to Basel I requirements to regularly and accurately report financial data have sounder banks. Based on data from 729 banks in 32 nations, Nier and Baumann (2006) find evidence that stronger market discipline induced by greater shares of uninsured liabilities and disclosure requirements boosts banks' capital positions, particularly in countries with more competitive banking systems.

Evidence on Bank Information Disclosure

Jordan, Peek, and Rosengren (2000) provide empirical evidence on bank disclosure by focusing on market reactions to bank supervisory actions. Market reactions are

stronger when banks failed to announce deteriorating conditions voluntarily than in the case of actual deteriorations. Thus, investors appear to punish banks for withholding relevant information. Jordan et al. find little evidence of contagion effects arising from announcements of supervisory actions, however.

Like any firm, banks have considerable discretion in their transmission of information. Landsman (2006) notes banks may be able to influence the information content of market signals via their timing of earning postings and balance-sheet adjustments. Karaoglu (2005), who examines various motivations for transfers of loans to third parties via loan sales and securitizations, finds evidence that U.S. bank holding companies use gains from loan transfers in part to influence both reported earnings and regulatory capital. Gunther and Moore (2003) in an examination of data from U.S. commercial banks between 1996 and 1998 find that banks in poorer financial condition are more likely to understate financial losses. They also conclude that the timing of supervisory examinations influences the accuracy of disclosures; disclosure of adverse results is more likely to occur in quarters in which supervisory examinations take place.

In a particularly creative study, Ashcraft and Bleakley (2006) use federal funds payments data to measure shocks to reserve balances, which in turn are used to identify the federal funds supply curve faced by an individual borrowing bank. Their analysis indicates first that funds suppliers do respond, albeit with a lag, to adverse changes in public information about a bank's credit quality. Then banks themselves respond by increasing leverage to offset future earnings effects, thereby managing the real information content of disclosures. Ashcraft and Bleakley conclude that public measures of a bank's loan portfolio performance provide information about future loan losses only in quarters when a bank is examined by regulators. The supply of funds does not respond during an examination quarter, however, which they suggest indicates that investors are unaware of information management by banks.

Does disclosure of information make banks less prone to failure? Tadesse (2005) utilizes World Bank survey data to develop an index measure of information disclosure for 49 countries covering 21 crisis episodes during much of the 1990s. He finds evidence suggesting that nations with both greater degrees of disclosure and more competitive banking systems are less likely to experience crisis episodes, which provides additional support for the Basel II disclosure requirements. In addition, Demirgüç-Kunt et al. (2008) examine Moody's financial strength ratings for more than 200 banks in 39 countries and find that banks in greater compliance with the information-disclosure requirements of Basel I exhibited greater soundness.

Thus, we have on the one hand considerable evidence that bond and equity investors can respond well enough to available information about banks' risk characteristics to induce movements in rates of return. There is much less evidence regarding the use of and responses to information on the part of uninsured depositors, but what we have indicates that depositors, like investors, use and react to the information available. There is also a good deal of evidence indicating that

government guarantees condition the responses of suppliers of funds to banks. Evidence also suggests that banks have the discretion to withhold certain information or to make income statement or balance sheet adjustments that influence the information content of the disclosures they do make. Finally, there is some evidence that a greater degree of information disclosure is associated with a less fragile banking system.

Evaluating the Market Discipline Pillar vis-à-vis the Other Pillars of Basel II

A key problem in evaluating the effects of the informational requirements of Basel II is that there is little research on whether bank managers actually respond to market signals. One study, by Bliss and Flannery (2000), examines the influence of the equity prices of U.S. bank holding companies on managerial actions, including adjustment of leverage, number of employees, and uninsured liabilities in relation to insured liabilities and total assets. Their conclusion is that variations in prices of bank stocks do not appear to influence regular managerial actions. The central market discipline element in the Basel II framework is disclosure of information relating to capital requirements. The market discipline pillar does not mandate full disclosure of all information. Instead, it establishes policies and timelines for disclosure of certain specific types of data.

The Limitations of Market Discipline under Basel II

The Basel II framework is in fact silent about other market discipline elements. It provides little explicit guidance about disclosures of other "material" information. Decisions about other disclosures deemed appropriate are generally left to bank directors and presumably varying levels of "moral suasion" or explicit requirements imposed by national regulators.

As Hamalainen, Hall, and Howcroft (2003) observe, Basel II provides no assurance that any debt holders will regard themselves at risk. It also specifies no provisions for creating market signals that are linked to regulatory actions. Arguably, the "market discipline" pillar is misnamed. Basel II does provide a foundation for satisfying a *necessary* condition for market discipline—namely, the disclosure of at least a body of information focused on a bank's balance sheet status in relation to its capital at particular points of time. There is little indication, however, that Basel II framers contemplated the *sufficient* conditions to enable markets to discipline via clear market signals and effective monitoring.

The Basel II market discipline pillar is only a first step toward providing a foundation for market discipline in nations that adopt the regulatory framework. Of course, for many nations this would be a significant first step. Barth et al. (2004,

2006) document that the absence of informational transparency is a key failing in many countries' banking systems. If, as they say, "supervision works best when it facilitates market monitoring," full implementation of the market discipline pillar of Basel II could result in dramatic improvements in banking systems around the globe.

Yet Basel II falls far short of the typical idea of market discipline in most of the academic literature. For banks in many developed nations, satisfying the market discipline pillar means little more than bringing measurement and reporting systems into line with statutory requirements. The actual information in the resulting disclosures is likely to be improved only in the sense that all banks will be reporting information based on similar methodologies ("internal ratings-based" measurement approaches for large banks, and a "standardized approach" for smaller banks). While uniformity of measurement methods and reporting should help funds suppliers in comparisons of institutions, and while new risk measures may be better than old measures and new reporting systems improve on old ones, it is not apparent that the flow of social benefits accruing to developed nations will necessarily exceed the social costs.

Theory versus Reality under Basel II's Market Discipline Pillar

Very little research has analyzed the interactions among the three types of pillars laid out in the Basel II framework. Llewellyn and Mayes (2003) find common sense conditions for market discipline and prompt corrective action by regulators to be complementary. Decamps, Rochet, and Roger (2004), use a continuous-time finance model to analyze implied market and book values of a representative bank with and without capital requirements, regulatory auditing, and market discipline. In the setting they contemplate, a representative bank receives one of two cash flows, depending on whether it incurs a cost to use a "good technology" to monitor its cash flow. Failing to incur this cost and opting for a "bad technology" invariably yields a negative cash flow. As long as the monitoring cost is relatively "low," there is a range of outcomes where the market value of using the bad technology exceeds than the market value of using the good technology, in which case a bank evades monitoring and increases its likelihood of failure. If the monitoring cost is high enough, bank shareholders voluntarily choose to shut down the bank.

In applying their model to Basel II's three pillars, Decamps et al. consider the regulatory imposition of a capital ("solvency") requirement, under which banks become illiquid before they become insolvent. This allows the regulator to determine when the cash flow is sufficiently low (or, equivalently, the monitoring cost is sufficiently high) that the bank will opt for the "bad technology." The regulator can then respond by closing the bank. If banks' cash flows are unobservable to

regulators without prohibitively expensive monitoring, then a higher capital ratio would be required to deter shirking.

Decamps et al. find that a subordinated debt payoff contingent on a bank's cash flow would reveal the cash flow and thereby save the regulator the cash-flow monitoring costs. Furthermore, the regulator can infer a cash flow below which it will audit the bank. Thus, mixing an audit policy, capital regulation, and monitoring of yields on subordinated debt generates an optimal Basel II-style three-pillar mix—provided that market securities prices are not so volatile that they fail to yield information about actual cash flows.

The conclusions reached by Decamps et al. hinge on two key assumptions. First, they assume regulators *require* banks to hold subordinated debt, which they show in reduces the bank's market value but allows for a lower capital requirement while inducing the bank to choose the good technology. Second, they assume regulators are free from political interference and follow *rules* for bank closures when cash flows and market signals from subordinated debt yields indicate imminent failures. Under these assumptions, the three pillars of capital regulation, market discipline, and a supervisory process can be reinforcing.

An attempt to move toward reinforcing these three pillars is certainly an improvement on Basel I, which focuses on capital regulation and pays no explicit attention to the other two pillars. Nevertheless, the Basel II framework does not require banks to issue minimum amounts of subordinated debt; nor does it provide for national regulators to monitor yields on bank debt issues. Rather, Basel II's implicit dismissal of rules in favor of discretion—and potentially politically motivated discretion—is the antithesis of the rules-based approach presumed in Decamps et al., which highlights its fundamental shortcoming. It pays insufficient attention to the useful regulatory role of market discipline and relies inappropriately on supervisory discretion—a key topic of the next chapter.

Summary: Market Discipline and the Banking Industry

- Market discipline entails reactions by suppliers of funds to changes in the risk composition of banks' balance sheets, either along the intensive margin via decisions by depositors, debt holders, or equity owners to reduce their supplies of funds or along the extensive margin by entirely cutting off their supply of funds to those banks. In order for the market discipline process is effective, suppliers of funds to banks must be able to engage in market monitoring so as to evaluate changes in banks' risk characteristics and financial conditions. To the extent that information derived from market-monitoring activities feeds back through direct market reactions to affect prices of secondary debt instruments issued by banks, corrective action on the part of banks may be generated.
- Several conditions must be satisfied for market discipline to induce corrective action at banks. Suppliers of funds to banks of correct information at appropriate

times, which the literature on information disclosure suggests is not necessarily always a likely outcome, even under legal disclosure mandates. In addition, suppliers' own funds must be at risk in the event of bank failure, so that they will have an incentive to react to perceived changes in the likelihood of bank insolvency in a manner that provides appropriate market signals. Furthermore, the market signals to a bank encountering a perceived change in its creditworthiness must be visible to all interested parties, including all funds suppliers. Finally, there must be incentives by an affected bank to respond to market signals in a solvency-promoting manner, by reducing risk exposures or replacing the bank's management.

- In light of the moral hazard problems associated with the provision of deposit insurance and the too-big-to-fail policy doctrine, most proposals for enhancing market discipline in banking have focused on mandated issuance of subordinated debts. Proposals for such mandates typically include features such as no-bailout clauses in subordinated-debt contracts, maturities of one to five years and relatively large denominations for such debts, staggered debt issue dates, and prompt corrective action rules such as debt covenants imposing sanctions in the event of adverse changes in a bank's performance.
- There has been little formal analysis of market discipline within theoretical models of the banking firm. A recent exception is the work of Landskroner and Paroush (2008), who propose the absolute elasticity of the rate of return on non-deposit liabilities with respect to the bank's beginning-of-period equity capital—that is, the proportionate change in the market rate of return on non-deposit liabilities in response to a proportionate change in the bank's equity—as a measure of the degree of market discipline imposed upon the bank. An elaboration of their theoretical framework indicates that the immediate effect of an exogenous increase in the value of this measure of market discipline is a reduction in the marginal cost of beginning-of-period equity capital, which raises the incentive for a bank to substitute away from deposit and non-deposit liabilities in favor of additional equity. Under certain conditions this action increases the likelihood that the bank attains positive gross interest income and thereby maintains a positive end-of-period value of equity.
- Although theoretical research suggests some degree of ambiguity regarding the effectiveness of such market signals in promoting disciplining effects, application of Landskroner and Paroush's analytical approach indicates that such effects are likely to diminish with decreased bank competition. In contrast, the effects of market discipline are likely to be greatest in highly competitive banking markets.
- Considerable empirical work suggests that rates of return on bank debts provide market signals that could promote market discipline. The evidence regarding whether regulators have a potential informational advantage over market processes is mixed. Although theory offers somewhat ambiguous predictions about whether market discipline and regulation are complementary or substitutable sources of corrective action, most empirical evidence suggests that regulatory actions tend to induce bank balance-sheet responses that at least partially undercut the strength of market discipline.

- The Basel II market discipline pillar's focus on promoting banks' timely release of more and better-quality information represents a step in the direction of truly promoting market discipline. This pillar otherwise does very little, however, in the way of laying a foundation for market discipline to yield corrective-action effects that are likely to promote significant improvements in bank safety and soundness.

Chapter 9
Regulation and the Structure of the Banking Industry

"[T]he real concern is supervision, not regulation. One needs to be sure that supervisors impose corrective measures or even close the bank before it is too late. The core of the problem is that any bank is always worth more alive than dead. This is so in particular because the informational capital of the bank is lost if it closes. So, even a competent and benevolent planner would always find preferable ex post to provide liquidity assistance to a bank in distress. But of course, if this is anticipated by bankers ex ante, this can be the source of moral hazard."

—Rochet (2008)

"[An] argument against a bank supervisory role for the Fed is the potential for moral hazard. To the extent that the Fed has institutional objectives other than maximizing social welfare, giving the central bank too broad a range of powers may invite abuse. For example, if the Fed were anxious to conceal the insolvency of some part of the banking system (an impulse that we have seen at times in other supervisory agencies), it might be tempted to distort interest rate policies in a way that increases bank profits or asset values, at the expense of macroeconomic objectives. Conversely, it is also possible that the Fed might use its supervisory authority to coerce banks into making loans that they otherwise would not make, in order to serve some goal such as providing short-term macroeconomic stimulus…"

—Bernanke (2001)

To this point, discussion of potential sources of interplay between bank market structure and regulation has emphasized how market structure influences the impacts of regulatory policies. This chapter focuses on regulatory and supervisory policies, particularly in the context of the Basel II standards, and how these policies themselves can impinge on the structure of banking markets. In addition, the chapter considers the implications of competition among bank regulators, a state of affairs that has been commonplace in the United States for some time and which likely will become more pervasive elsewhere in the coming years.

D. Van Hoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-642-02821-2_9, © Springer-Verlag Berlin Heidelberg 2010

Public Interest versus Public Choice Perspectives on Bank Regulation

Regulation of the banking industry is so omnipresent that most observers take for granted the regulatory superstructures that exist in many countries. What factors motivate national governments to regulate banks? Is bank regulation in the public interest, or could other motives be at work?

Public Interest and the Alleged "Need" for Bank Regulation

Basic economics principles textbooks, such as Miller (2009) go to great pains to point that the word "need" is objectively undefinable. "Need," the new student learns, is a word that economists learn to avoid. People have wants and desires, and economists can seek to tease from the data the relative values that individuals place on these wants and desires. Thus, the notion of "needs," the textbooks suggest, has no place in discussions of economic issues.

Nevertheless, banking researchers and policymakers commonly allege a "need" for bank regulation. Santos (2001, p. 46), for instance, notes that "[t]he justification for any regulation usually stems from a market failure such as externalities, market power, or asymmetry of information between buyers and sellers." He suggests that "[i]n the case of banking, there is still no consensus on whether banks need to be regulated. . ." Vives (2008, p. 445) argues that "[t]he need for regulation is particular acute when charter values are low, such that incentives to take risks are high, and the social cost of failure is high—making it so that banking failure has a large impact." Carletti (2008) contends that "[t]he potential instability of the banking system and the need [for] consumer protection are the fundamental rationales behind the introduction and development of regulation."

Perhaps this notion that there must always be a "need" for regulation in order for any regulation to exist helps to explain why a number of modern financial economists insist that any theoretical analysis of bank regulation makes sense only in a setting in which regulation of the banking industry has the potential to be unambiguously welfare-improving. That is, any legitimate framework of analysis for examining the implications of regulation of the banking industry must include a *public interest rationale* for regulation. Otherwise, why would government entities ever choose to regulate banks?

Public Choice Motivations for Bank Regulation

In fact, it has long been well understood in the industrial organization literature that there are motivations for regulating firms that have absolutely nothing to do with aiming to correct market failures. Certainly, the existence of third-party spillover costs or benefits, inefficiencies arising from market power, or welfare losses owing

to informational asymmetries can provide a rationale for policy actions to correct any or all such market failures. Other rationales for regulation exist, however. Officials in government agencies possessing the power to regulate may, as suggested by Posner (1971), desire to marshal public resources to transfer economic rents from one group to another group, the latter of which may be in political favor or may offer to provide implicit or explicit rewards to officials in return for supervising such rent transfers. Or, as proposed by Stigler (1971), firms may seek out regulation from government agencies, with an aim to "capturing" regulatory officials who will have the power to protect firms from competition from prospective new entrants.

In point of fact, financial economists who insist that analysis of bank regulation makes sense only in the context of models in which regulation is "needed" are out of tune with the field of regulatory economics. In their survey of this area, Viscusi et al. (2005, pp. 376–378) review the traditional public interest theory of regulation, which they refer to as "normative analysis as positive theory." Viscusi et al. note that this theory puts forth the hypothesis that regulation "occurs when it should occur because the potential for a net social welfare gain generates public demand for regulation"—that is, a perceived regulatory "need." They contend, however, that "[l]acking in this analysis is a description of the mechanism that allows the public to bring this result about."

According to Viscusi et al., "the key reason why [normative analysis as positive theory] has lacked supporters. . .is the large amount of evidence that refutes it," because "many industries have been regulated for which there is no efficiency rationale" and "in many cases, firms supported or even lobbied for regulation." Viscusi et al. also contend that even a common reformulation of theory of regulation purely in the public interest—that regulation initially is established to correct a market failure but then often is mismanaged by the agency charged with performing such regulation—still fails to square with the evidence. That is, as originally argued by Posner (1974), "theoretical and empirical research. . .[has] demonstrated that regulation is not positively correlated with the presence of external economies or diseconomies or with monopolistic market structure."

Applying the Economic Theory of Regulation to the Banking Industry

Figure 9.1 depicts application of the economic theory of regulation developed by Peltzman (1976) to banking in the case in which banks face no fixed costs and constant marginal resource costs. In this situation, as discussed in Chapter 3, point C in panel (b) is the perfectly competitive equilibrium, and point M is a monopoly outcome, with points between C and M representing alternative cases of imperfect competition. The inverted-U-shaped curve in panel (a) graphs profits against the market loan rate. Firms earn zero economic profits at point C in panel (b) and maximum profits at point M.

The preferences of the regulator depend on industry profits and on the loan rate. Higher profits are associated with healthier—and more satisfied—firms and hence yield positive utility to the regulator. A lower loan rate implies an increased level of credit at better terms, which brings about approval by consumers, nonfinancial firms,

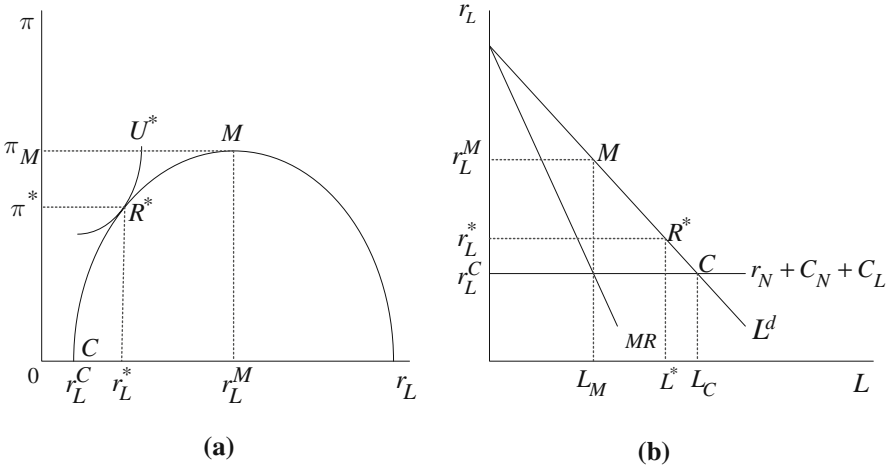

Fig. 9.1 A regulatory optimum

and elected representatives and thereby also generates utility for the regulator. Thus, the regulator's indifference curves slope upward, as depicted in panel (a), with the regulator's utility increasing as indifference curves shift upward and to the left. In panel (a), the highest feasible indifference level of utility attainable by the regulator, given the associated market demand and the cost structure in panel (b), is U^*. Point R^*, therefore, is the regulator's optimal outcome.

In the situation depicted in panel (a) of Fig. 9.1, the regulator desires a banking industry structure roughly midway between the extremes of perfect competition and monopoly. If an unregulated industry is already in equilibrium at a point near R^*, there is little incentive for either interest group to strive for establishment of a bank regulatory regime. If an unregulated banking industry is presently in an equilibrium at or near point M, however, consumers, nonfinancial firms, and elected representatives will have an incentive to press for industry regulation, because they perceive that point R^* will be closer to their preferred point C at which the loan rate is lowest loan rate and the level of credit greatest. Alternatively, if an unregulated banking industry is presently in an equilibrium at or near point C, the industry itself has an incentive to lobby for regulation, because R^* is closer to banks' preferred, maximum-profit point M. Clearly, in either case, one or the other set of interest groups will have more to gain from regulation, and interest group competition is likely to become fierce, as in Becker (1983).

Assessing the Implications of the Economic Theory of Regulation

The economic theory of regulation suggests that once a bank regulatory regime is put into place, actual profit and loan rate (and hence aggregate credit) outcomes will depend on the preferences of the regulator. Consider panel (a) of Fig. 9.2. The indifference curves for this banking regulator are relatively steep, implying that the

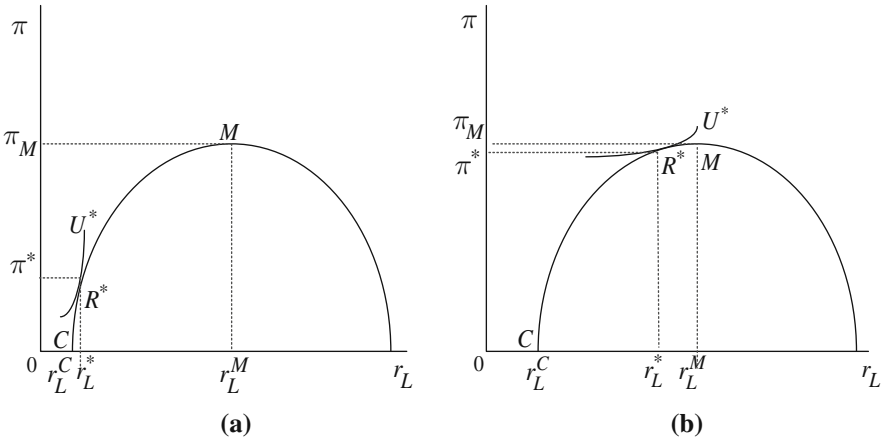

Fig. 9.2 Differences in regulator preferences

regulator is willing to accept significant swings in bank profits in exchange for slight changes in the loan rate. As a consequence, the regulatory optimum R^* must be relatively close to the competitive limit. In contrast, in panel (b), the regulator desires to maintain industry profits within a fairly narrow range and hence is willing to permit wide variations in the loan rate (and hence aggregate credit), implying relatively shallow indifference curves. The resulting regulatory optimum is nearer to the monopoly limit.

Ceteris paribus, natural interpretations of the regulatory optima depicted in Fig. 9.2 are that the case in panel (a) applies to a bank regulator that is relatively attuned to the public interest, whereas the situation in panel (b) characterizes a regulator that is more nearly captured by the industry along lines suggested by Stigler (1971). Rosenbluth and Schaap (2003) provide evidence from 22 industrialized nations suggesting that differences in electoral structures help to determine which type of bank regulatory regime is more likely to result, and Stiroh and Strahan (2003) offer analyses of U.S. deregulation between 1976 and 1994 that would seem to suggest a movement from a real-world situation similar to panel (b) to one more consistent with panel (a). Masciandaro and Quintyn (2008) argue that a penchant for both concentrated banking markets and consolidation of power within fewer regulatory bodies suggests greater likelihood of capture, and based on data from 88 nations, they find evidence supporting both predispositions.

As discussed in Chapter 6, however, there some theories and a limited body of empirical evidence suggesting that risks of banking failure may rise with increased competition. Thus, a regulator selecting an optimum such as that depicted in panel (b) might actually be aiming to pursue the public interest by assuring bank profitability that helps limit the scope for failures. Alternatively, in light of the abundance of theories and mixed evidence regarding the relationship between bank competition and risk, a regulator exhibiting a choice such as depicted in panel (b) could offer risk reduction as a rationale for that choice in an effort to camouflage the fact that the regulator actually has been captured by the industry.

Indeed, Hardy (2006) argues that banking regulators are particularly susceptible to industry capture. He notes that, consistent with Laffont and Tirole's (1991) criteria for industries prone to regulatory capture, the banking industry is concentrated, has much at stake, and involves several forms of informational asymmetries, while financial regulations typically are complex from the perspective of an outsider. In such a setting, an individual bank can gain from market stability, which means that the industry as a whole has an incentive to seek out some forms of supervision, preferably from malleable supervisors willing to protect established banks from additional competition. Hardy suggests, however, that capture of banking regulators has the potential to result in relatively more benign outcomes, in which regulators may face incentives to impose costly and constraining supervision. In fact, Hardy concludes, it is possible that regulations that emerge are costlier and more constraining than consistent with maximized social welfare.

A Generalized Perspective on Evaluating Bank Regulation

The economic theory of regulation implies that in principle, regulation could arise if market failures are relatively unimportant characteristics of banking markets. Certainly, the existence of large external spillovers, significant market power, or widespread informational asymmetries in banking markets help to provide motivation for establishment of banking regulatory regimes, such factors are not required to motivate a perceived "need" for government action. Indeed, taking the stance that market failures must be a feature of any banking model aimed at analyzing effects of regulation can be regarded as a means to attaining a self-fulfilling prediction from the model that regulation is indeed "needed." After all, if a market failure is present in a theory, then the theory naturally will suggest that the failure can be corrected via an appropriate regime of taxation or regulation, even if the reality is that regulation is driven by other motivations.

To be sure, there are good reasons to posit that sources of market failures exist in the banking industry. As discussed in Chapter 4, network externalities arise through bank payment-system interactions. The analyses of Chapters 3 through 5 suggest that market power could well be of importance in banking. Furthermore, asymmetric-information issues abound in financial markets. Nevertheless, the significance of these potential market failures is an empirical issue. Previous chapters indicate that to date, evidence regarding the magnitudes of these various potential sources of market failures is mixed.

Regarding the "consumer protection" rationale for regulation noted by Carletti (2008), which has been offered as motivation for a number of policies aimed at assuring "fair" financial service prices (see, for instance, Hannan, 2007), and the like, Benston (2000) correctly notes that most nations already have broad social regulations to address such issues. Consequently, bank-specific regulations aimed at protecting consumers are unnecessary. Furthermore, a number of consumer-protection regulations, such as the Community Reinvestment Act in the United States, amount to requirements for banks to engage in cross-subsidization activities.

Economic theory suggests that governments seeking outcomes requiring subsidization of certain consumers could achieve them more efficiently by directly subsidizing those consumers rather than requiring banks to cross subsidize.

Irrespective of the empirical significance of the other, market-failure-motivated rationales for banking regulation, there is considerable evidence favoring adopting a broader, public-choice perspective. Abrams and Settle (1993), for instance, provide evidence that U.S. banks were able to use regulation to shape the structure of the industry during the 1930s, and Kroszner and Strahan (1999, 2001) reach a similar conclusion for the deregulation period of the 1970s and 1980s. In addition, Lown and Wood (2003) conclude that over time U.S. banks have been able to induce regulators to prescribe required reserve ratios consistent with levels desired by the industry. More than market failures explain the historical and contemporaneous patterns of banking regulation.

The Political Economy of Banking Supervision Conducted by Multiple Regulators: Is a "Race to the Bottom" Unavoidable?

Of course, banks do face risky prospects and can be susceptible to runs. Banks also face adverse selection and moral hazard problems. Some of these market failure are both inherent to the markets in which they operate, which in some contexts truly do provide rationales for contemplating and implementing certain forms of regulation. Nevertheless, as Winston (2006) emphasizes, poorly designed or improperly managed regulatory frameworks can deliver social welfare losses that rival or even exceed those of market failures that the policy frameworks were created to mitigate.

Within some nations, such as the United States, and some regional trading areas, such as the European Union, overlapping jurisdictions already are commonplace, and in a number of locales banks can effectively *choose* their supervisors. As a consequence, it is conceivable that bank supervisors themselves face incentives to engage in a form of regulatory competition that might, in principle, result in a "race to the bottom" in terms of supervisory standards, thereby laying a foundation for a higher rate of bank failures.

Regulatory Preferences and Bank Closure Policies

The public choice perspective on regulation of industry emphasizes that the preferences of regulators are crucial in shaping the structure of an industry supervisory regime. Before considering interactions among regulators with common jurisdictions, it is important to understand how a regulator's preferences influence its supervisory actions.

There is a relatively small literature on how a regulator's preferences shape its actions when confronted with situations in which they must choose between engaging in forbearance—permitting a bank that has violated *ex ante* supervisory policies

to continue to operate via ex post, discretionary non-enforcement of such policies versus closing the bank (see Bhattacharya et al., 1998, for a good overview). Many studies focus on the issues faced by a cost-minimizing deposit insurer/supervisor. For instance, Acharya and Dreyfus (1989) derive optimal bank-closure policies for a regulator committed to minimizing the deposit insurer's liability, in which regulatory forbearance is always suboptimal. Allen and Saunders (1993) examine how the put-option value of deposit insurance to a bank may be over-estimated if a regulator opts for forbearance for threatened banks that reduce their risk exposures. Dreyfus et al. (1994) argue that under some circumstances an insurer/regulator may prefer a higher limit on its exposure to deposit risks if a bank's larger deposit base enables it to continue making payments to holders of non-deposit liabilities, thereby making a regulatory forbearance policy more cost-effective. Fries et al. (1997) contemplate an environment in which bankruptcy costs and incentives for bank equity holders to recapitalize troubled institutions respond endogenously to regulatory policies and conclude that forbearance can sometimes be an optimal policy.

Such analyses, however, apply only to settings in which a regulator's credibility and reputation are not at issue. In contrast to Campbell et al.'s (1992) assumption that regulators are averse to monitoring effort, Boot and Thakor (1993) examine a two-period game-theoretic model in which the private payoff of the regulator depends on his monitoring reputation. A regulator charged with monitoring and closing banks can either respond to a sub-optimal first-period risk choice by requesting a better risk choice the next period or closing the bank. They find that the equilibrium policy that emerges from this principal (society)-agent (regulator) problem is socially sub-optimal, because imperfectly informed agents view a bank's closure as a signal that the regulator may have poorly monitored the closed institution. Thus, in an effort to avoid developing a reputation as a poor monitor, a regulator engages in more forbearance than society otherwise would prefer. They conclude that society would gain from leaving the decision regarding closure to an a regulatory agent not directly involved in supervisory monitoring, redoubling efforts to prevent industry capture of bank regulators, placing strict limits on the allowable scope of forbearance on the part of regulators, or limiting the riskiness of bank asset choices to make it easier for society to judge the quality of supervisory monitoring by regulators.

Mailath and Mester (1994) also undertake a game-theoretic analysis of bank closure policies. They focus on a trade-off between the influence of closure policies on bank risk-taking and a social cost of closure resulting in the loss of intermediation services to society. They find that in such a setting commitment to a closure policy typically is infeasible and in some circumstances is neither cost-minimizing for a deposit insurer nor optimal for society as a whole. That is, forbearance emerges in many circumstances as the cost-minimizing and/or socially optimal policy. Mailath and Mester also show that as bank size increases, forbearance is more likely, suggesting that banks have an incentive to become "too big to fail." Mailath and Mester note that removing regulatory discretion would make closure threats more credible but at the cost of potentially closing individual banks that might otherwise recover, hence reducing ex post social welfare in such circumstances. Therefore,

"[i]f the economy is expected to improve, then giving the regulator discretion may be beneficial" (Mailath and Mester, 1994, p. 293).

Competition among Bank Regulators

Do discretionary regulators with overlapping jurisdictions face incentives to engage supervisory-standards competitions that can result in minimal standards and hence an increased probability that regulators will face a closure-versus-forbearance choice? Surprisingly little work has been done to address this important question, in spite of the fact that jurisdictional overlap has long existed in the United States, where a commercial bank can effectively switch regulators by changing its national charter to a state charter, or *vice versa*—known as a "charter flip"—or by altering its Federal Reserve System membership status. Whalen (2002), for instance, finds that key factors explaining charter flips are the competitive structures of banking market, variations in banks' riskiness, and the levels of ratings received from supervisors. Notably, Whalen finds that charter flips are associated with lower ratings from an initial regulator, suggesting that banks do sometimes attempt to "shop around" for less demanding regulators. In addition, Whalen (2008) finds evidence that higher supervisory costs associated with a national bank charter makes prospective U.S. bank entrants more likely to opt for a state charter when they are contemplating entering relatively concentrated markets.

A Theory of Optimal Supervisory Choices of a Single Bank Regulator

Weinberg (2002) provides a very intuitive model of supervisory choices by one or more bank regulators with overlapping jurisdictions. In the model, a regulator's key task is to choose two variables—the probability of a bank examination (denoted p in Fig. 9.3) aimed at revealing a bank's choice between high risk that leads to failure and low risk that does not and a fee (denoted f in the figure) charged to banks that do not undertake risky actions that lead to failure. The regulator makes this choice to maximize the expected net income of a typical bank less examination costs incurred by the regulator, which in Weinberg's model amounts to minimizing examination costs. If the regulator incurs the cost of conducting an examination and finds that the bank has made a high-risk choice, the bank is closed at a cost to the deposit insurance system, in which case the bank receives no income.

In this setting, Weinberg shows that the opportunities set of p-f combinations available to the regulator is the shaded area bounded by two constraints in Fig. 9.3, which is drawn under two assumptions: (i) Examination costs on average are less than net income of non-failing banks; and (2) any bank failures are covered out of the regulator's budget, necessitating collecting fees to cover such failure costs even if the regulator conducts no supervisory examinations. One constraint is the regulator's budget constraint, B. Given the regulator's resources, the set of budget-feasible choices lies on or below and to the right of this budget constraint. The other constraint is the incentive-compatibility constraint, IC. On or above and to the left

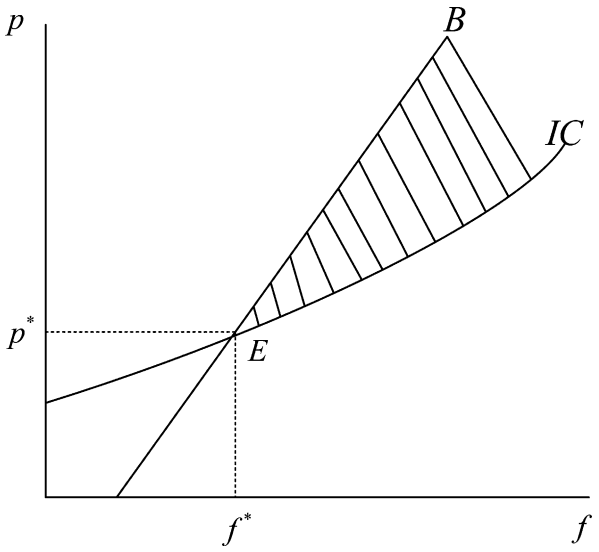

Fig. 9.3 A single regulator's supervisory policy choice

of this constraint, the examination fee is sufficiently low and the probability of being examined sufficiently great that the bank will choose the less-risky action.

The efficient policy choice is the one with lowest feasible probability of examining a bank sufficient to satisfy the incentive compatibility constraint, which is the choice p^* at point E, with the budget constraint implying an associated fee f^*. In the face of a greater incentive problem faced by banks that causes the *IC* constraint to shift upward, there would be a movement to a new efficient point above and to the right of point E. That is, the regulator would optimally respond by increasing the probability of an examination and raising the examination fee.

Figure 9.4 displays Weinberg's supervisory-choice analysis for a single regulator when one set of banks is inherently riskier than another, with the innately riskier banks implying the incentive compatibility constraint IC'; that is, *ceteris paribus*, more frequent examinations are necessary to induce the inherently higher-risk banks to take less-risky actions consistent with avoiding failures. For inherently less-risky banks, the incentive compatibility constraint *IC* applies. Separate regulation of the two groups would entail more frequent, higher-fee supervisory examinations for the innately more risky group.

A regulator cannot distinguish between the two groups ex ante—that is, between examinations—so it must set a common *p-f* pairing for all banks. If the regulator knows that the inherently more risky group constitutes a small share of the banks it supervises, then one option is to choose to establish a policy configuration that focuses on inducing the much larger number of innately less-risky banks to make less risky decisions. If the regulator takes for this approach, then it opt for a budget consistent with a reduced the frequency of examinations relative to the case in which

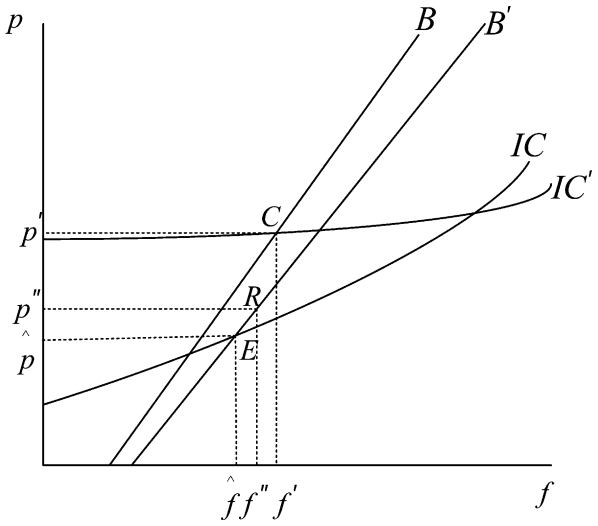

Fig. 9.4 Two types of banks

all banks are less risky, yielding the budget constraint *B*' in Fig. 9.4. Point *E* depicts the efficient policy configuration, $\hat{p}$ and $\hat{f}$ selected by a regulator that selects this approach.

Of course, such a choice implies a regulatory preference function consistent with allowing a bank failure to occur from time to time between examinations. A regulator that prefers an even lower bank failure rate will prefer to induce *all* banks, including the innately more risky banks, to make choices that will contribute to lower risk of failure. Such a "conservative" regulator will opt for the *p-f* pairing given by p' and f' at point *C* in Fig. 9.4. Hence, the more conservative regulator will select a greater probability of examinations and higher examination fee than would a regulator that optimizes by selecting point *E*.

The Case of Competing Regulators

To apply Fig. 9.4 to the case of competing regulators confronting inherently high- and low-risk banks, suppose that a conservative regulator opts for the policy configuration at point *C*. A somewhat less conservative, competing regulator, however, is satisfied experiencing isolated bank failure per unit of time and hence will be more willing to select a configuration such as p'' and f'' at point *R*. Naturally, if the latter regulator has a desire for a relatively larger "clientele" of banks to supervise, this gives it an even greater incentive to select the supervisory configuration *R*.

Indeed, other things being equal, all banks would flock from regulator selecting configuration *C* to the less conservative regulator opting for point *R*, perhaps by engaging in charter flips. The more conservative regulator would be able to maintain some share of the overall industry clientele only if it could offer some other

inducement, such as services not available to those offered by the regulator selecting the policy configuration R. Absent such an inducement, the more conservative regulator's clientele would diminish over time. Alternatively, at some point the more conservative regulator might be replaced by one more willing to experience rare failures, resulting in a competition leading to point R—or perhaps ultimately point E, at which all regulators are willing to observe a low but steady rate of failures, an outcome that Rosen (2003) suggests characterized the U.S. experience between 1983 and the late 1990s. Naturally, the actual outcome would hinge on those appointing the regulators, which under Basel II consists of regulators possessing whatever degree of discretion is granted by national governments.

A Supervisory Race to the Bottom?

As discussed in Chapter 6, from the mid-1990s through the late 2000s, U.S. deposit insurance essentially amounted to a taxpayer guarantee. By and large, client banks only paid fees to compensate their regulators—the Federal Reserve, the Federal Deposit Insurance Corporation (FDIC), and the Office of the Comptroller of the Currency (OCC)—for variable and fixed expenses associated with conducting periodic examinations. The amount of deposit insurance reserves maintained by the FDIC was so small that the failure of two moderate-sized institutions in 2008 wiped out the bulk of the reserves, requiring the FDIC to obtain assistance from taxpayers via recourse to the U.S. Treasury.

In Weinberg's framework, Fig. 9.5 applies when bank failures are funded outside regulators' budgets—an environment he refers to as one of unconsolidated budget constraints. Because regulators collect fees from banks only to cover the expenses associated with examinations, in this setting the regulator's budget constraint is a ray from the origin. Budget constraint B applies in the case in which all banks are inherently equally risky, as in Fig. 9.3. Analogously to the analysis in Fig. 9.3, the policy configuration p^* and f^* at point A is the efficient outcome. This point is also analogous to point C in Fig. 9.4; a conservative regulator can rest assured that choosing point A and maintaining a relatively "active" examination schedule will be the efficient policy if a small set of banks becomes more innately risky.

Now consider the case in which such a small group of innately risky banks exists. As in Fig. 9.4, the budget constraint B' is applicable for a regulator that is willing to accept a positive failure rate. The regulator satisfied with a positive failure rate can specify the same fee, f^*, within a smaller budget with a lower probability of examinations $\tilde{p}$, at point Q, than will a conservative regulator that otherwise would select point A. In Weinberg's model, innately more risky banks will prefer the policy configuration at point Q because it enables them to take riskier actions between examinations than allowed by the configuration at point A. Thus, if the regulator with budget constraint B' has sufficient preference for expanding its clientele—or perhaps simply has a strong preference for a "quiet life" involving fewer supervisory examinations—it will opt for the configuration at point Q. As a consequence, banks will begin to flip charters and switch from the regulator selecting point A in favor of the regulator opting for point Q.

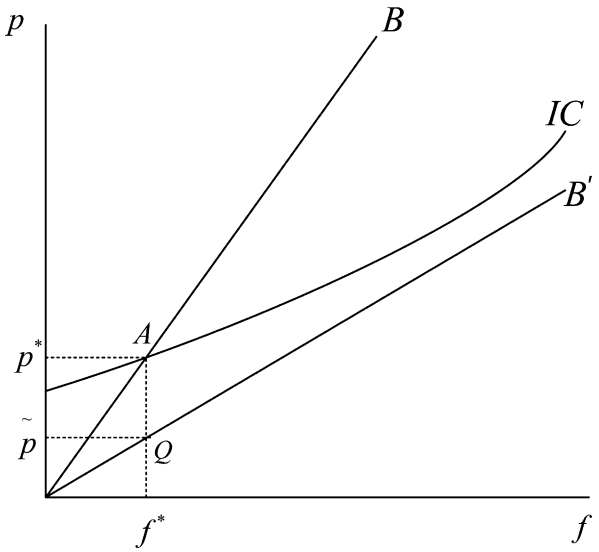

Fig. 9.5 The case of unconsolidated regulatory budget constraints

If the conservative regulator either rearranges its preference weighting in favor of retaining its clientele or is replaced during the passage of time by one willing to be less conservative, then it can bid to maintain client banks by reducing its fees. Indeed, a regulatory equilibrium at the origin, with $p = f = 0$, represents a feasible outcome of a bidding war involving sufficiently non-conservative regulators. Thus, a true regulatory race to the bottom can take place. Weinberg's analysis suggests that a race to the bottom in a competition among is avoidable. Avoiding regulatory failure requires national regulators to operate with consolidated budgets and with sufficient degrees of conservatism regarding the magnitude of the potential bank failure rate.

Applying Weinberg's analysis indicates that the environment confronted by U.S. regulators in the 2000s was not inconsistent with a regulatory race *toward* the bottom. The budgets of the Federal Reserve and the OCC are unconsolidated, and the zero-premium FDIC policy in conjunction with woefully insufficient reserves implied an essentially unconsolidated budget as well for the FDIC.

Should Bank Regulation Be in the Hands of Monetary Policymakers?

Peek et al. (2003, 2001, 1999) document evidence that they argue favors the view that synergies exist between bank regulation and monetary policymaking. The Federal Reserve's proprietary information regarding the status of non-market-traded banks that it regulates, they contend, gives it an advantage over other agents in

forecasting inflation and unemployment, thereby providing a rationale for activist monetary policy.

Nevertheless, in the context of a theoretical model in which a conservative political party is less willing to countenance inflation than a liberal party, Franck and Krausz (2008) find that the conservative party will prefer, as long as the likelihood of a banking crisis is sufficiently low, for banking supervision to be operated independently from the central bank, which in their model is consistent with lower inflation. Some support for the hypothesis that housing bank regulation within a central bank is inflationary is provided by Di Noia and Di Giorgio (1999), who examine the experiences of 23 OECD nations, of which about half conduct all banking regulation and supervisory tasks within central banks. They find that in nations in which central banks are the sole regulators, average inflation is higher. Additionally, they conclude that loan-rate markups above and deposit-rate markdowns beneath government security rates are higher, consistent with Cukierman's (1992, pp. 117–129) argument that a central bank interest in supporting banking markets via interest-rate stabilization the can translate into inflation-boosting increases in the monetary base.

Goodhart (2000) and Goodhart and Schoenmaker (1995a, b) evaluate the cases for and against placing bank regulatory functions within a central bank. They conclude that, on one hand, combining monetary policymaking and regulatory functions under a single umbrella is useful in the coordination of rescues of struggling banks. On the other hand, because modern economies' financial systems now encompass many more institutions than traditional banks, such coordination gains from combining monetary-policy and regulatory functions within a central bank are shrinking. Furthermore, in broadened financial systems, in the event of a crisis, taxpayers likely face losses that extend beyond the banking industry, giving them an incentive to establish a separate bank regulatory authority. Finally, there is a potential for conflicts of interest to arise in conducting both policy tasks, potentially including higher inflation, consistent with Bernanke's (2001) suggestion that a central bank such as the Federal Reserve may also feel pressure to maintain interest rates at levels inconsistent with optimal macroeconomic policy. In addition, Bernanke notes, a central bank might have an incentive to press banks it regulates to expand lending to help stimulate the economy. Perhaps in light of such concerns, Vives (2001) proposes restructuring the euro-area bank regulatory framework in such a way that gives the European Central Bank access to supervisory information but places actual supervision in the hands of a separate regulatory agency.

Ioannidou (2005) examines 1990–1998 U.S. data regarding monetary policy actions on the part of the Federal Reserve and formal regulatory actions undertaken by the Federal Reserve, the FDIC, and the OCC. Ioannidou finds evidence that the Federal Reserve monetary policymaking and bank regulatory supervision are intertwined. He finds that the Federal Reserve loosens its supervisory stance after implementing contractionary monetary policy actions but concludes that its supervisory actions do not feed back to influence its monetary policymaking. Supervisory positions of the FDIC and OCC were unaffected by Federal Reserve monetary policy actions. In contrast, Čihák and Podpiera (2008), who utilize data on bank

supervisory structures and quality of supervision from 84 countries during the early 2000s to investigate whether greater supervisory integration is associated with higher-quality supervision, find no evidence that housing supervision within or outside of central banks generates any significant differences in supervisory quality. Clearly, the evidence regarding the desirability of a dual role for central banks as monetary policymakers and bank regulators is mixed, although Weinberg's analysis suggests possible concerns in light of the unconsolidated nature of the Federal Reserve's budget vis-à-vis costs of closing failed banks—an issue not addressed in the rest of the literature on this topic.

The Supervisory Review Process Pillar of Basel II

The overlapping regulatory jurisdictions of the Federal Reserve, the FDIC, and the OCC, in which a charter flip or Federal Reserve membership change can alter the regulatory authority supervising a bank, are unique to the United States. Nevertheless, in Europe and other parts of a world with increasingly internationalized banking markets, regulatory jurisdictions are becoming progressively more overlapped.

Undoubtedly, recognition of this fact helps to explain why national governments have been working toward bring their regulatory systems into line with the supervisory review process pillar of the Basel II framework. There are four key principles of Basel II's supervisory review process. The first is

> Principle 1: Banks should have a process for assessing their overall capital adequacy in relation to their risk profile and a strategy for maintaining their capital levels (p. 205)

The document goes on to state that "supervisors will typically require (or encourage) banks to operate with a buffer, over and above the Pillar 1 standard. . .," and that "for banks to fall below minimum regulatory capital requirements is a serious matter that may place banks in breach of the relevant law and/or prompt non-discretionary corrective action on the part of supervisors."

Clearly, this principle is related to Pillar 1 capital requirements. It gives national regulators considerable discretion determining whether banks are well, adequately, or inadequately capitalized in relation to the minimum ratios specified by the Basel II agreement.

The next principle again focuses on Pillar 1:

> Principle 2: Supervisors should review and evaluate banks' internal capital adequacy assessments and strategies, as well as their ability to monitor and ensure their compliance with regulatory capital ratios. Supervisors should take appropriate supervisory action if they are not satisfied with results of this process (p. 209)

Left unstated is how and when supervisors should undertake a review of banks' internal assessments and strategies. In a supplementary two-page discussion of Principle 2, we read only about "periodic review" that may involve on- or off-site examinations or discussions, external auditing, and "periodic reporting." The document also is silent as to the "appropriate supervisory action" to undertake if national regulators are unsatisfied after a review.
Principle 3 also focuses attention back to Pillar 1:

> Principle 3: Supervisors should expect banks to operate above the minimum regulatory capital ratios and should have the ability to require banks to hold capital in excess of the minimum (p. 211).

A brief discussion of Principle 3 notes that "supervisors will need to consider whether the particular features of the markets for which they are responsible are adequately covered," and that "supervisors will typically require (or encourage) banks to operate with a buffer, over and above the Pillar 1 standard." In addition, it suggests:

> There are several means available to supervisors for ensuring that individual banks are operating with adequate levels of capital. Among other methods, the supervisor may set trigger and target capital ratios or define categories above minimum ratios (e.g., well capitalized and adequately capitalized) for identifying the capitalization level of the bank.

The final principle is as follows:

> Principle 4: Supervisors should seek to intervene at an early stage to prevent capital from falling below the minimum levels required to support the risk characteristics of a particular bank and should require rapid remedial action if capital is not maintained or restored (p. 212).

The BIS document's detailed discussion of Principle 4 is brief enough to be quoted in full:

> Supervisors should consider a range of options if they become concerned that a bank is not meeting the requirements embodied in the supervisory principles outlined above. These actions may include intensifying the monitoring of the bank, restricting the payment of dividends, requiring the bank to prepare and implement a satisfactory capital adequacy restoration plan, and requiring the bank to raise additional capital immediately. Supervisors should have the discretion to use the tools suited to the circumstances of the bank and its operating environment.
>
> The permanent solution to banks' difficulties is not always increased capital. However, some of the required measures (such as improving systems and controls) may take a period of time to implement. Therefore, increased capital might be used as an interim measure while

> permanent measures to improve the bank's position are being put in place. Once these permanent measures have been put in place and have been seen by supervisors to be effective, the interim increase in capital requirements can be removed.

Principles 3 and 4 clearly allow considerable regulatory flexibility in the setting of minimum capital standards. Taken together, they give a national regulator the discretion to pursue responses ranging from encouraging to requiring a bank to abide by the Pillar 1 capital standards. It is also clear that under Basel II regulators have the discretion to give a bank considerable time to adjust capital positions in the event of inadequacy.

The Supervisory Review Process Pillar: Conceptual Issues

Corrective action, of course, falls under the bank supervisory review process, and many observers are highly critical of the wide scope for regulatory discretion granted by Basel II. For instance, Hamalainen, Hall, and Howcroft (2003) critique the Basel II standards for failing to include provisions aimed at avoiding regulatory forbearance. In addition, the Shadow Financial Regulatory Committee in "Statement on the Basel Committee's Revised Capital Accord Proposal" (2001) that:

> Increased discretion for banks and regulators would likely result in increased opportunities for risk arbitrage by banks and greater potential for regulatory forbearance, both of which undermine effective capital regulation. Regulatory evaluations of bank risk and capital requirements would differ across banks within the same country and across countries, depending on bank choices and differences in the latitude regulators in particular countries grant banks. The number, complexity, and opaqueness of the new rules established under the Basel proposal would add to regulatory forbearance by making it harder to hold regulators accountable for their judgments about bank risk. It is worth noting that American and British regulators currently do not agree even about the appropriate method to measure the probability of loan default using historical data. Given that absence of agreement, the potential for regulatory inconsistency is great.

Several assumptions are implicit in criticisms such as these. One, of course, is that discretionary policymaking is inferior to policymaking based on a system of rules. A second assumption is that rapid corrective action is typically more likely to avoid social losses than a gradualist approach. A third is that international coordination of bank regulation is desirable. Let's contemplate each assumption in turn.

Discretion Versus Rules

A traditional assumption is that regulation induces management actions that the less-precise disciplining effects of markets cannot. At the same time, there are inherent difficulties in government regulation of financial institutions. Regulation can be inefficient. It sometimes can create perverse outcomes, such as greater moral hazard problems. Regulators can be slow to adjust to institutional and market innovations.

They may not be able to adapt to complexities of large and multifaceted financial institutions. This can lead to regulatory forbearance in the face of suspected weaknesses at troubled institutions.

Are regulatory actions triggered by particular events preferred over discretion in bank regulation? There is now a considerable literature weighing rules versus discretion in economic policymaking (see Mishkin (2006) for a useful discussion with regard to bank regulation). The key message of this literature is that a policymaker with discretion faces a time-inconsistency problem: An announcement by a policymaker that it will take a certain future action ultimately has little force if agents recognize that the discretionary policymaker in fact is willing to diverge from that intention. In the monetary policy literature, this recognition has produced a near-consensus view that society benefits from making monetary policymaking independent from political influence and appointing policymakers who are innately more conservative in the sense of placing more loss weight on reducing inflation.

Analogously, a society focused on the overall policy objective of a safe and sound banking industry are more likely to succeed by developing a mechanism for commitment, typically an institutional structure involving clear *ex ante* supervisory rules that banks know will be followed ex post. This fundamental conclusion, of course, rules out discretion in bank regulatory policy and is consistent with the literature in favor of rules such as those discussed by Boot and Thakor (1993).

How Tough Should a Supervisory Policy Rule Really Be?

The appropriate form and speed of actions specified in regulatory rules is nonetheless a subject of continuing debate. As noted by Rochet (2008), a regulator would always prefer ex post to keep from losing the informational value offered by a bank and hence has an incentive to provide assistance to a distressed bank. Because bank managers realize this, regulators confront a moral hazard problem in designing rules for closing banks.

The prompt corrective action provisions of the 1991 FDIC Improvement Act provide a set of tripwires leading up to implementation of a well defined bank closure process (see John Walter, 2004 for a description of this process). Freixas and Parigi (2007) develop a theoretical analysis of U.S. prompt corrective action rules and find that they mirror an optimal policy forthcoming from their model, in which banks face increasingly tough limitations on their holdings of risky assets as the degree of capital adequacy declines.

Sleet and Smith (2000) use a two-period banking model to examine a setting with deposit insurance, a discount window, and imperfect ability to distinguish between liquidity versus solvency problems. They suggest that promptly liquidating the assets of troubled banks may not always be appropriate, even when insolvency appears relatively likely. This conclusion arises naturally enough given their assumption that there are social costs associated with bank liquidations. In a setting in which entrepreneurs trade off the gains from expropriating funds acquired from bank loans against lost collateral, Kocherlakota and Shim (2007) find that optimal

regulation entails forbearance if a collapse of collateral value is relatively unlikely, ex ante, but requires prompt liquidation of the assets of problem banks otherwise.

Mishkin (1999) contends that "financial consolidation is now moving toward a banking system with a smaller number of large," nationwide, diversified banks... that "are less likely to fail," so that "deposit insurance is no longer as needed." While acknowledging that regulators must be vigilant, he concludes that the prompt corrective action provisions of FDICIA increased the likelihood that "supervisors will do their job properly and prevent excessive risk taking on the part of banks." Mishkin also suggests addressing the moral hazard problem associated with the too-big-to-fail policy via a regulatory stance of "constructive ambiguity." Under this policy, in the event of a systemic event affecting numerous banks, regulators would stand ready to permit the first large bank facing insolvency to fail. Thus, Mishkin argues, depositors and creditors of each too-large-to fail bank would have an incentive to discipline managers that take on significant risks and expose them to potential losses.

Shim (2006) applies a dynamic financial model of bank-regulator interaction, capital regulation, private information about returns, costly liquidation, and capacity to hide risks, and argues that a policy of randomized bank closures or bailouts is preferable to a policy of prompt corrective action without a bailout option. (This conclusion may be at least partly contingent on Shim's assumption of risk-based deposit insurance.) Thus, to date, the theoretical literature suggests that prompt action to close banks is not necessarily the optimal supervisory policy.

Of course, weighed against such theoretical ambiguities are the practical experiences of regulators in many countries during the 1980s and 1990s and again since 2006. In the case of the U.S. savings institutions crisis of the 1980s, for instance, the sounding of alarms by Kane (1985) and others was answered by regulatory forbearance rather than corrective action. The result, as documented by Kane (1989), Barth (1991), and White (1991), was a huge loss to the U.S. deposit insurance system.

This experience prompted enactment of the FDIC Improvement Act (FDICIA) of 1991, which requires accountability in the form of "tripwires" for escalating disciplinary actions on the part of regulators. While this is not a completely rules-based prompt corrective action process, it does establish a set of benchmarks that are aimed at constraining regulatory discretion. Unfortunately, in the early stages of the most recent financial crisis, the Federal Reserve misinterpreted the banking impacts of deflating real estate- and mortgage-market bubbles—to which its own too-loose monetary policy had contributed—as a liquidity crunch. The Federal Reserve proceeded to transform itself from a lender of last resort into the lender of first resort, thereby preempting application of prompt corrective action rules. As demonstrated by Taylor and Williams (2009) and discussed in a broader context by Taylor (2009), the Federal Reserve's establishment of lending programs for banks actually had virtually no impact on aggregate liquidity. The credit it extended did, however, keep the real estate- and mortgage-market bubbles inflated for more than a year before they ultimately collapsed, creating a massive solvency crisis for over-extended banks.

In the context of the analysis of Mailath and Mester (1994), the Federal Reserve acted in a discretionary manner with regard to troubled institutions whose loss

it worried might be costly to society, under an anticipation that the economy would improve once the Federal Reserve provided more liquidity. If so, this turned out to be an incorrect anticipation. Or perhaps, as in the analysis of Boot and Thakor (1993), Federal Reserve officials were driven by a worry over lost reputation. Alternatively, perhaps the Federal Reserve was simply a captured regulator. Whatever the Federal Reserve's motivation, it is arguable that if FDICIA prompt corrective action rules had been applied to the banks experiencing declines in credit quality at that early stage of the crisis, a signal would have been sent to other banks to cut back on subprime and other lower-quality loans. Instead, the Federal Reserve's decision to extend too-big-to-fail reasoning to moderately large banks at the outset created a massive moral hazard problem by signaling to all but the obviously sufficiently-small-to-fail institutions that the Federal Reserve would not permit them to fail. Most post mortems of the recent financial downturn indicate that many—perhaps the majority—of the worst-quality loans were extended between mid-2007 and mid-2008. In retrospect, therefore, it is clear that proper application of prompt-corrective-action tripwires could have done much to reduce the scope of the subprime meltdown.

When Is International Coordination of Bank Regulation Appropriate?

Relatively little work examines whether international coordination of regulatory supervisory rules is truly appropriate. One exception is Bliss (2008), who emphasizes potential problems associated with jurisdictional disputes among national regulators confronting insolvencies of large, multinational financial holding companies. Resulting disputes among regulators, he suggests, can contribute to failure to compel sufficiently speedy remedial actions.

Another exception is Holthausen and Rønde (2005), who develop a theoretical framework to analyze the behavior of regulators when a multinational bank operates in two countries. A host-country regulator provides information about the bank's activities to the home-country regulator. If the interests of the regulators do not coincide, the host-country regulator may fail to provide full information to the home-country regulator, and either the bank either is not closed when it should be, or it is closed when it should not be. When regulators have divergent interests, social welfare can be improved by establishment of a multinational regulator—that is, a system of fully coordinating national regulators much like the approach in the European Union—to make closure decisions.

Finally, Dell'Ariccia and Marquez (2006) also examine a two-country model. They assume that, as a result of structural interdependence between the two nations' banking systems, externalities induce regulators to choose suboptimally low standards. In their model, centralizing regulation by making a supranational authority responsible for setting standards internalizes these externalities and eliminates this problem. There is a cost, however, which is reduced flexibility to adjust to cross-country banking differences. Dell'Ariccia and Marquez find that optimally trading

off the benefits and the costs of centralization implies that centralized regulation is more likely to be socially optimal for nations with relatively homogeneous banking systems. They also conclude that obtaining the social optimum via supranational regulation requires tougher supervisory standards than would have been established by the national regulator with the highest standards among all regulators.

Is There Really a Basel II Supervisory Review Process?

Basel II's process for supervisory review allows for considerable variation in regulatory standards. The theoretical literature previously discussed suggests that under certain circumstances this approach might be appropriate. But if heterogeneities across banking systems are so pronounced as to give wide latitude for discretion in enforcement of standards, theory would also suggest that internationally coordinated standards may not be appropriate.

Specifying a wide range of discretion for national bank regulators in fact obviates any true process for national governments adopting the Basel II framework. The guidelines for the supervisory review pillar essentially can be summed in one single sentence in the document: "*Supervisors should have the discretion to use the tools suited to the circumstances of the bank and its operating environment.*" Ultimately, Basel II does not require national regulators to specify clear rules that will guide their supervision procedures. It also is silent regarding rules governing enforcement actions against poorly managed and troubled institutions. Thus, if owners and managers of banks under Basel II will operate under internationally coordinated standards, the standards effectively will be enforced in whatever way national supervisors see fit to enforce them, often in situations involving regulatory conflicts, as discussed in detail by Wall and Eisenbeis (2000). Rather than ensuring that Basel II establishes supervisory standards at least as tough as those of the most stringent national regulator, its framers have opted to allow each participating nation the discretion to race to the bottom of the range of standards.

The supervisory process pillar is unambiguously the weakest of the three Basel II pillars. It is unlikely to do anything to promote increased bank safety and soundness. Kaufman (2006) is probably correct in concluding that a U.S.-style prompt corrective action rule is a preferred alternative over the supervisory process pillar of Basel II.

Indeed, Barth et al. (2004, 2006) provide strong evidence that provision of considerable scope for regulatory discretion is likely to be counterproductive. They suggest that nations that have granted greater discretion to bank supervisors have tended to have banking systems that exhibit less development, more corruption, and poorer overall operating performance. To the extent that the supervisory process pillar sanctions greater discretion on the part of national regulators, it actually could prove to be detrimental to global banking development and stability. Indeed, Barth et al. (2008) conclude that if anything, regulatory changes in a number of countries have weakened the likelihood that supervision will promote stability. In a study of loan terms established by 278 commercial banks in 39 nations to borrowers in 83

countries, Magalhaes and Tribó (2009) provide some support for this conclusion by finding that other things being equal, a higher level of supervisory authority is associated with riskier and less diversified bank lending.

Furthermore, in a study of bank productivity across 22 nations between 1999 and 2006, Delis et al. (2008) find neither Basel II supervisory process regulations nor capital requirements contribute positively to banking productivity. In contrast to conclusions reached by Tirtiroğlu et al. (2005) in a study of U.S. banking data, Delis et al. find that certain regulations restricting bank operations tend to enhance bank productivity—notably those aimed at promoting market discipline. Pasiouras et al. (2008) likewise conclude based on data from 615 banks spanning 74 countries in the early 2000s that market-discipline-enhancing regulations enhance profit efficiency. Pasiouras et al. find, however, that restrictions on bank activities depress cost efficiency while enhancing profit efficiency.

Regulatory Compliance Costs and Industry Structure

Complying with banking regulations is also a costly activity, as documented by Demirgüç-Kunt et al. (2004). Based on a study of more than 1,400 banks across 72 nations, Demirgüç-Kunt et al. find that supervisory regulation has significantly contributed to a higher cost of financial intermediation. To explore the scope of regulatory costs faced by banks, let's begin by considering the compliance-cost implications of the Basel II standards.

Assessing Banks' Costs of Basel II Compliance: Economies of Regulation?

As presently conceived by U.S. regulators, the Basel II framework for bank regulation will entail a segmented system of risk-based capital requirements for banking institutions. In one category will be perhaps as many as a dozen banks with more than $250 billion in assets or at least $10 billion in on-balance-sheet foreign risk exposures. These specific institutions will be required to adopt a so-called advanced internal ratings-based approach, referred to henceforth as a "full Basel II" requirements system. In addition, other banks meeting minimal infrastructure requirements can voluntarily adopt the full Basel II framework. Remaining institutions will, depending on exactly how U.S. regulators ultimately decide to phase in Basel II rules, either continue to be subject to essentially the rules previously in place under Basel I or opt to be regulated under a "Basel IA" framework utilizing an expanded set of risk weightings for required capital, particularly in relation to mortgage loans. Originally, small and medium-sized institutions were to have been regulated under the Basel IA, but regulators presently are discussing an alternative, as-yet-undefined alternative "standardized approach," perhaps modeled on the European methodology, to capital risk assessments.

Surprisingly, in spite of the fact that virtually all regulators and bankers agree that expenditures are likely to be relatively high, there has been very little investigation of this question. When the Basel II framework is implemented, how many dollars' worth of resources will U.S. banks expend in order to comply?

Rule-of-Thumb Estimates

Most references to the costs of complying with Basel II refer to rule-of-thumb estimates. For instance, it is common to see references in the financial media to speculation that 10 percent or more of banks' information-technology spending has been allocated to Basel II compliance since the early 2000s. Another rule of thumb (see, for instance, Hitt et al., 1998) is that as much as 15 percent of banks' non-interest expenditures are information-technology-related.

Such rule-of-thumb estimates typically do not separate out quasi-fixed setup costs from ongoing compliance costs. As a first approximation, however, suppose that these percentages apply to the present value of all combined expenses related to Basel II compliance, and consider applying such rules of thumb to aggregate U.S. banking data implies that in 2006, when the FDIC (Quarterly Banking Profile, 2006) reports that non-interest expenses for the 7,402 commercial banks then operating of about $290.3 billion. Conservatively assuming that 10 percent of these were costs related to information technology and that 10 percent of these expenditures were related to Basel II compliance, the implication is that the estimated total cost Basel II compliance is roughly $2.9 billion. This averages out to total compliance costs of just over $390,000 per U.S. commercial bank.

This estimate is very imprecise. Suppose, for instance, that banks' information-technology-related expenses really are as much as 15 percent of their non-interest expenditures. Then applying the 10 percent rule of thumb for Basel II compliance costs as a share of information technology spending to the 2006 data boosts the estimated overall Basel II compliance cost—again, conservatively assumed to equal the discounted present value of the entire stream of Basel II-related costs—to nearly $4.4 billion (or almost $590,000 per commercial bank). Naturally, reducing either or both of the rule-of-thumb percentages cuts into the resulting point estimate. For example, if the share of banks' non-interest expenditures devoted to information technology is actually only 5 percent and the share of the latter expenses related to Basel II compliance is only 5 percent, then the point estimate for Basel II costs drops to only about $725 million (about $98,000 per bank).

Estimates of Basel II Compliance Costs Based on Survey Data

Instead of relying on rules of thumb, the Office of the Comptroller of the Currency (2006a, b) has surveyed banks about their own estimates of costs in implementing with the new Basel II rules. In principle, the OCC's approach offers hope for somewhat greater precision in estimating compliance costs.

The OCC concludes that for the 65 *nationally chartered* banks it judges to be most likely to utilize either Basel II or Basel IA requirements, the combined compliance costs (discounted present value for 2006) amount to about $473 million. Across these 65 commercial banks, this estimate implies an average total compliance cost of just under $7.3 million per institution. The OCC further estimates that if *all* nationally chartered banks were to adopt Basel II or Basel IA regulatory requirements, the combined compliance costs (again as a discounted present value of all costs as of 2006) would be nearly $1.1 billion, or almost $680,000 per institution.

Of course, these OCC estimates apply only to nationally chartered banks, or just over 20 percent of the U.S. banking system (albeit a portion containing many of the largest U.S. banks). Thus, the low-end total compliance-cost estimate of $725 million yielded by a rule-of-thumb-style approach is almost certainly too low. Indeed, these OCC cost estimates for nationally chartered banks make the overall compliance-cost range of $2.9 billion to $4.4 billion begin to look more reasonable.

Based on banks' survey responses, the OCC found that among commercial banks required only to meet Basel IA standards, an average institution with assets of less than $100 million faced a total (discounted present value in 2006) cost of $100,000; a bank with assets between $100 million and $1 billion confronted a total cost of $500,000; a bank with assets between $1 billion and $10 billion faced a total cost of $1 million; and a bank with assets exceeding $10 billion confronted a total cost of $3 million. Finally, the OCC estimated that a typical bank utilizing the full Basel II regulatory framework faced a total compliance cost of about $21 million.

Now consider the result of applying these OCC compliance-cost estimates to the entire U.S. banking system. In 2006, the FDIC (Quarterly Banking Profile) reports that there were 3,246 banks with assets less than $100 million, 3,662 banks with assets between $100 million and $1 billion, 406 banks with assets between $1 billion and $10 billion, and 88 banks with assets exceeding $10 billion. Of the last set, let's suppose for the sake of conservatism that the OCC's 19 banks are the only banks that initiate full implementation of Basel II regulatory standards and that the remaining 69 very large banks continue to utilize Basel IA standards. Applying the OCC's per-bank estimates within each size class and totaling yields a total compliance cost estimate of almost $2.8 billion, or close to the $2.9 billion rule-of-thumb estimate obtained above.

"Economies-of-Basel II"—Scale Advantages in Basel II Compliance?

The OCC's (2006a, b) estimates of total compliance costs by bank size can be utilized to develop rough estimates of average costs per dollar of assets of implementing either the standardized Basel 1A or full Basel II systems. Suppose that we use the midpoint of each classification discussed above for the Basel IA banks (for instance, $50 million for banks with assets less than $100 million, $550 million for banks with assets between $100 million and $1 billion, and so on) as the denominator of average-compliance-cost ratios for banks with assets falling within the ranges with endpoints. According to *American Banker* (2007), on December 31, 2006 the average assets of a bank among the top 25 commercial banks in the United States (among which presumably is the set required to use or opting to utilize the full Basel

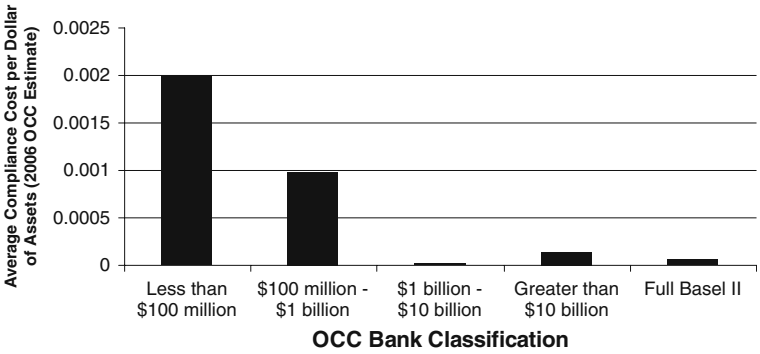

Fig. 9.6 Average compliance costs by asset category

II framework, denoted "Full Basel II" in the figure) was about $339 billion. Among all others exceeding $10 billion in assets, the average bank had about $21 billion in assets. To obtain rough estimates of average compliance costs in relation to assets, these figures were used as denominators for the largest two groups of banks.

The resulting estimates of average compliance costs in relation to bank assets, displayed in Fig. 9.6, are suggestive of a relationship that yields two potential implications. First, there appears to be a likelihood of significant diseconomies of scale in terms of Basel II compliance costs for banks with less than $1 billion in assets. On a proportionate basis, the OCC's estimates indicate that implementation costs may well be significantly higher for smaller banks than for larger banks.

Second, among banks with more than $1 billion in assets, there may be compliance-cost diseconomies of scale for Basel IA banks that increase their scale from the $1 billion-$10 billion size category to the greater than $10 billion size classification. Indeed, the OCC's estimates indicate that at some point beyond an asset size of $10 billion, a bank experiences lower average compliance costs by opting for advanced internal ratings-based regulation under the full Basel II framework

Figure 9.6 hints at the potential for significant average compliance-cost differentials that might well emerge if a more thorough-going statistical study could be conducted utilizing more precise data on compliance costs over the full range of the U.S. bank-asset-size distribution. Is it possible to generalize about costs specific to Basel II versus other non-interest expenses? On one hand, the OCC (2006b) notes that the 19 nationally chartered banks that are engaged in full Basel II implementation reported a per-bank average of total information technology expenses of $42 million that could be related to Basel II implementation. But these banks guessed that roughly half of those expenditures involved adoption of information technologies that would have been put into place in any event, hence the OCC's $21 million estimate of per-bank compliance costs. Undoubtedly, this "guesstimate" lies behind commonly encountered estimates that as much as 50 percent of banks' expenses on information technologies may be Basel II-related. On the other hand, this "guesstimate" only applies to banks implementing the full Basel II system. The OCC's compliance-cost estimates for smaller institutions suggest that banks adopting Basel IA surely incur compliance costs that are much less significant shares of their budgets for information technologies and other business lines.

In fact, it may well be that actual compliance costs for the U.S. banking system will turn out to be closer to or even to exceed the higher-end rule-of-thumb estimate of $4.4 billion. Recent informal statements by some bankers indicate that full Basel II compliance costs have turned out to be at least 50 percent higher than originally forecast and reported in the media. In addition, some surveys (see, for instance, Accenture, 2005 and Risk Transfer Magazine, 2005) suggest that Basel II compliance cost projections for a number of internationally active European banks increased as these banks, which advanced more rapidly toward Basel II compliance than their U.S. counterparts, encountered previously unanticipated implementation costs. Sloan (2007), for instance, reports that internationally active banks with more than about $135 billion in assets now anticipate spending an average of about $67 million on Basel II compliance and that 20 percent of those banks now estimate compliance costs nearly double that amount. Furthermore, a recent survey by Ernst and Young (2006) suggests a global average full Basel II compliance cost of $70 billion but also indicates that the top five Canadian banks' implementation costs exceed this amount.

Note that if these latter figures are correct, then the OCC estimate of a full Basel II compliance cost of $21 million per institution is at least three times too low, implying less of a scale advantage from full Basel II implementation than indicated in Fig. 9.6. Indeed, assuming that the OCC's Basel IA compliance-cost estimates are not also understated, a full Basel II implementation cost of $67 million would boost the average compliance cost for a large bank required to utilize the full Basel II approach to a level about 40 percent *higher* than the average cost experienced by other large banks opting to the standardized Basel 1A regulatory methodology. In this case, the back-of-the-envelope calculations involved in producing the last bar displayed in Fig. 9.6 would no longer hold true. The smallest of the very large banks required to comply with the full Basel II methodology would encounter a proportionate compliance-cost *dis*advantage relative to somewhat smaller institutions with greater than $10 billion in assets that are able to utilize the standardized Basel 1A methodology. Furthermore, the institutions within the latter group obviously would not have a compliance-cost incentive to opt for full Basel II treatment after all.

To What Extent Do Basel II Compliance Costs "Matter"?

What are the implications of Basel II compliance costs for the U.S. banking system? The first part of the answer to this question relies on total discounted-present-value estimates of Basel II implementation costs. The second part of the answer depends on an assessment of the on-going variable costs associated with complying with the new regulatory framework.

Total Implementation Costs

To begin with, what does a discounted-present-value estimate of total Basel II implementation costs for the U.S. banking system of, say, $2.8 billion or $4.4, billion really mean? Based on data for the entire U.S. banking system in 2006, total

compliance-cost estimates within the range of these amounts imply an average Basel II implementation expense ranging from just under 1 cent to perhaps about 1.5 cents per dollar of non-interest expenses. In relation to the aggregate 2006 net income of the U.S. banking system of just over $128.6 billion, these total compliance cost estimates would imply the equivalent of a one-time 2–3.5 percent average charge against net income—a relatively significant "regulatory tax" on bank owners. Of course, if the recent reports of higher-than-anticipated compliance costs have any validity, the aggregate tax on bank profitability may well be higher.

Differences across individual institutions in the total costs of implementing Basel II regulations stand to influence the distributional effects they will have on the banking system. Such cost differences obviously matter for large institutions that must choose between the standardized Basel 1A methodology versus the full Basel II internal-ratings-based approach. As discussed above, however, there are uncertainties regarding the eventual scope of the costs of implementing the full Basel II system. Hence, it is presently unclear whether a situation along the lines of the right-hand side of Fig. 9.6 exists, or, alternatively, whether full Basel II implementation costs in fact are turning out to be much higher, so that the opposite may be true. One thing that is certain is that in light of this uncertainty, some large banks facing a choice regarding whether to opt for the standardized approach or the internal-ratings-based approach have a difficult decision to make.

As noted above, it also appears likely that average costs of small community banks are likely to be disproportionately increased during the transition to the standardized Basel 1A framework. In the near term, smaller banks contemplating choices with respect to mergers or acquisitions might find that lower compliance costs associated with increased scale could be a factor that pushes them to act.

From an economic standpoint, there may not be much more to be discerned from analysis of discounted-present-value estimates of all anticipated costs to be incurred by banks in implementing Basel II. Estimates of discounted present values of streams of costs are of assistance in strategic choices of strategic *paths* to pursue—such as whether to choose between Basel IA or Basel II procedures or to opt for a change in scale in an effort to contain average costs.

Variable Costs and On-Going Effects of Basel II

Once a bank and its competitors have chosen a strategic path to pursue, the factors that drive their behavior and, hence, influence market outcomes are the determinants of *flows* of revenues, expenses, and profits. Thus, effects of Basel II on banks' *variable* costs are what impinge on banks' week-to-week and month-to-month balance-sheet decisions—their lending and other portfolio-allocation choices and their decisions regarding liability mix—as well as their off-balance-sheet activities—securitization, derivatives trading, and so on.

Unfortunately, very little appears to be known about the on-going effects that compliance with the Basel II framework are likely to have on banks' variable costs. Indeed, only a single point estimate appears to be available. As part of the its evaluation of the compliance costs faced by a typical bank operating under the

internal-ratings-based approach of the full Basel II system, the OCC provides an estimate of an annual expense of $2.4 million.

It is perhaps unsurprising that estimates of the flows of variable costs associated with Basel II implementation are nearly nonexistent. After all, economists struggle to measure the relevant variable costs that banks already face under the existing regulatory framework. Evaluating the on-going burden of Basel II compliance clearly will require collection of variable-cost data once operations under Basel II begin.

Implementation of Basel II likely will entail on- and off-balance-sheet adjustments entailing a number of potential benefits and costs for the banking system and for the economy as a whole. In the context of a general-equilibrium calibrated simulation, the most conservative of Van den Heuvel's (2008) estimates of the on-going costs to society generated by capital-regulation-induced liquidity reductions is 0.1 percent of total consumption, or, based on 2006 data and in 2006 dollars, an annual decrease in consumption somewhat below $10 billion. Thus, while Basel II compliance costs are not trivial, in is conceivable that they are significantly less than the broader costs associated with bank balance sheet adjustment brought about by the imposition of capital requirements. Furthermore, *both* compliance costs *and* costs related to broader on- and off-balance-sheet adjustments—which Tchana (2007) suggests society will deem worth incurring if shocks are large enough and members of society are sufficiently risk-average—together will ultimately determine the full competitive ramifications of Basel II adoption for the banking system.

Bank Regulation and Endogenous Fixed Costs

As discussed in Chapter 4, recent work by Dick (2006, 2007) applies Sutton's (1991) theory of endogenous sunk fixed costs to the banking industry and evaluates the consumer welfare effects of quality-influencing mergers. Her research indicates that as long as scope for active quality competition remains in place, consumer welfare will not necessarily be harmed by a merger that maintains or adds to the degree of market concentration. Essentially, Dick's work suggests that the banking industry may be a sort of "natural oligopoly" with negligible implications for social welfare.

Regulatory Compliance Costs: A Missing Component?

Nevertheless, there are good reasons to exercise caution in assessing applications of Sutton's theory for the banking industry. Sutton's own work (1991, 2001) has applied his theory to advertising- and technology-intensive industries in which endogenous sunk costs are important. While many such industries face a number of social regulations, such as product quality regulations, most arguably face nothing like the broad range and depth of economic regulations confronted by the banking industry. Indeed, costs imposed on banks by regulations constitute a substantial component of the fixed outlays that banks incur. In a review of regulatory compliance costs imposed on banks, Elliehausen (1998) concludes that the total cost of complying with U.S. bank regulations as of the early 1990s amounted to about

12–13 percent of banks' noninterest expenses. Although Elliehausen suggests that labor costs are the major component of startup and ongoing costs of complying with banking regulations, the fact that banks must comply suggests that an important consequence of regulation is to transform a significant portion of labor costs that otherwise would be variable into fixed costs. In fact, Elliehausen argues about half of the required activities of bank employees relating to regulatory compliance are undertaken *only* because they are required, suggesting that at least half of labor expenses related to regulatory compliance are purely regulatory fixed costs.

The fixed-cost burden of bank regulation undoubtedly has increased as a consequence of implementation of the Basel I capital standards beginning in the early 1990s and the gradual phase-in of the more highly capital-intensive Basel II regulations since the mid-2000s. As noted above, the total Basel II compliance cost faced by the banking industry may be equivalent to a charge against a single year's earnings of as much as 3.5 percent, with a significant component of this cost likely being fixed costs associated with utilization of information technologies required to implement Basel II requirements.

It is arguable that fixed costs of regulation should be regarded as exogenous sunk costs. After all, even though there are several regulators—the Federal Reserve, the Office of the Comptroller of the Currency, the Federal Deposit Insurance Corporation, the Office of Thrift Supervision, the National Credit Union Administration, and relevant state authorities—every federally insured U.S. banking institution confronts the same regulatory superstructure and faces essentially similar compliance-cost-creating layers of supervision and regulation. Surely a portion of these regulatory compliance costs are essentially exogenously fixed across institutions.

Regulatory Sunk Fixed Costs

The fact that estimated average costs of complying with Basel II differ considerably across larger and smaller banks, a significant component of regulatory costs surely is endogenous. Furthermore, endogeneity of regulatory fixed costs is a prediction that naturally flows from the economic theories of regulation discussed earlier in this chapter. Over time, changes in external factors—such as a subprime mortgage crisis—are likely to alter the terms of the trade-off between the interests of regulated firms and consumers, resulting in a shift in the nature of the regulatory equilibrium—and hence the magnitudes of fixed costs faced by banks in complying with the shifting terms of regulation that they confront. Banks have considerable input into the process by which regulations are adopted and adapted in response to external events. Hence, a significant portion of the fixed regulatory compliance costs they face arguably are endogenous.

Direct application of Sutton's theory to the banking industry presumes that privately incurred fixed bank expenses—such as those associated with advertising or, as in Dick's (2007) work, branch and ATM networks and other outlays aimed at enhancing product quality—are the primary components of banks' endogenous sunk fixed costs. Available evidence suggests, however, that the majority of regulatory

compliance costs are also sunk fixed costs that comprise a considerable portion of banks' noninterest expenses. Work to date on applying Sutton's framework to banking ignores this important source of endogenous sunk fixed costs in the banking industry. Of course, integrating such costs into a Sutton-style theory would require accounting for endogeneity of the process by which a regulatory equilibrium is attained and maintained over time, which likely would be a challenging undertaking.

Regulation has altered, currently impinges on, and will continue to affect significantly the history, present status, and likely future configuration of the banking industry. A key component of banks' sunk fixed costs is compliance costs generated by regulation, which by their nature are both faendogenous and subject to influence by banks themselves. Consequently, to the extent that endogenous sunk fixed costs help to explain the present structure of banking markets, one fundamental reason may well be that an endogenous regulatory framework, rather than a natural, unregulated process may be at least partly—or perhaps even largely—responsible. Until the likely role of endogenous sunk fixed *regulatory* costs in determining the equilibrium structure of the banking industry has resolved, policymakers should not accept at face value the notion that banking may be regarded as essentially a type of "natural oligopoly."

Summary: Regulation and Bank Industry Structure

- Academics and policymakers sometimes refer to a "need" for bank regulation. Certainly, there are persuasive theoretical rationales for regulation purely in the public interest aimed at addressing potential market failures such as monopoly power, informational asymmetries, or externalities that may exist in banking markets. Nevertheless, whether public interest rationales are sufficient to justify regulation is an empirical issue. The literature on economic regulation has long recognized that governments may choose to regulate markets for reasons that have nothing to do with market failures. Indeed, applying the theory of regulation to banking suggests public choice rationales for regulation, including the extreme situation of regulatory capture, in which the regulator seeks to satisfy only the preferences of firms in the regulated industry. Thus, a presumption by some banking researchers that any theoretical framework purporting to analyze the effects of banking regulation must incorporate a correctable market failure is misplaced and arguably biased in favor of promoting regulatory activism.
- Theories of optimal bank closure policies indicate that regulatory forbearance can be welfare-improving under some circumstances, with the incentive for a regulator to engage in forbearance rising as bank size increases. Nevertheless, application of the economic theory of regulation to settings in which regulators compete for regulated bank "clienteles" demonstrates that a "race to the bottom" in terms of regulatory standards and socially unwarranted forbearance can result. A key condition contributing to such an outcome is when at least one competing regulator has an unconsolidated budget, meaning that it does not have to draw

from its own budget to cover costs created by failed institutions. Conflicts of interest between monetary and banking policy responsibilities of central banks can potentially make them poor choices to serve as bank regulators, but the theory and evidence on this issue is mixed.

- Some studies of optimal supervisory bank policies argue in favor of discretion over rules, but prompt-corrective-action rules receive support from the economic literature that seeks to evaluate social benefits of policy rules rather than discretion. Relatively little research has been devoted to studying appropriate formulation of international supervisory policies in the banking arena. So far work in this are suggests that in environments in which international interactions are important, the socially optimal policy is either to assign bank supervisory decisions to a supranational regulator or to base supervisory policies on those pursued by the national regulator with the highest standards. The Basel II supervisory process pillar invests banking supervision with national authorities, to which it grants considerable discretion in establishing supervisory standards. Thus, this pillar eschews rule-based policymaking and establishes conditions that arguably could favor an international race to the bottom in bank supervision.
- The compliance costs that banks will have to incur to satisfy the Basel II regulatory framework are poorly understood. Based on available estimates, however, these costs likely amount to the equivalent of average annual earnings of banks between 2.5 and 3.5 percent. These costs, which on a per-asset-dollar basis appear to be higher for smaller banks, would be added to the significant costs that banks already incur in complying with the full range of bank regulations, which have been estimated to be 12 to 13 percent of banks' total noninterest expenses.
- Taken together, regulatory compliance costs account for a significant portion of the fixed costs of operating banking institutions. To the extent that banks can help frame the structure of regulation, therefore, regulatory compliance costs arguably are a component of endogenous sunk fixed costs in banking. To date, application of the theory of endogenous sunk fixed costs in determination of bank market structure has failed to take into account this potential regulation-based explanation for the relatively concentrated structure of the banking industry.

References

Abrams, Burton, and Russell Settle, 1993, Pressure-group influence and institutional change: Branch-banking legislation during the Great Depression, Public Choice 77, 687–705.

Accenture, 2005, Banks slow in making progress toward Basel compliance, Accenture survey shows, July 25 (http://accenture.tekgroup.com/article_print.cfm?)

Acharya, Sankarshan, and Jean-Frncois Dreyfus, 1989, Optimal bank reorganization policies and the pricing of federal deposit insurance, Journal of Finance 44, 1313–1333.

Acharya, Viral, 2009, A theory of systemic risk and design of prudential bank regulation, Journal of Financial Stability, 5, 224–255.

Acharya, Viral, Iftekhar Hasan, and Anthony Saunders, 2006, Should banks be diversified? Evidence from individual bank loan portfolios, Journal of Business 79, 1355–1412.

Acharya, Viral, and Tanju Yorulmazer, 2008, Information contagion and bank herding, Journal of Money, Credit, and Banking 40, 215–231.

Acharya, Viral, and Tanju Yorulmazer, 2007, Too many to fail—An analysis of time-inconsistency in bank closure policies, Journal of Financial Intermediation 16, 1–31.

Adams, Robert, and Dean Amel, 2007, The effects of past entry, market consolidation, and expansion by incumbents on the probability of entry, Finance and Economics Discussion Series No. 2007-51, Board of Governors of the Federal Reserve System.

Adams, Robert, Lars-Hedrik Röller, and Robin Sickles, 2002, Market power in outputs and inputs: An empirical application to banking, Manuscript, Board of Governors of the Federal Reserve System, October 23.

Agarwal, Sumit, and Robert Hauswald, 2007, Distance and informational asymmetries in lending, Manuscript, Federal Reserve Bank of Chicago and American University, March 1.

Aggarwal, Raj, and Kevin Jacques, 1997, A simultaneous equations estimation of the impact of prompt corrective action on bank capital and risk, In Financial Services at the Crossroads: Capital Regulation in the 21st Century, Federal Reserve Bank of New York Conference Proceedings, February 26–27.

Agoraki, Maria-Eleni, Manthos Delis, and Fotios Pasiouras, 2008, Regulations, competition, and bank risk-taking in transition countries, University of Bath Working Paper 2008.09.

Akhavein, Jalal, Allen Berger, and David Humphrey, 1997, The effects of megamergers on efficiency and prices: Evidence from a bank profit function, Review of Industrial Organization 12, 95–139.

Akhavein, Jalal, Lawrence Goldberg, and Lawrence White, 2004, Small banks, small business, and relationships: An empirical study of lending to small firms, Journal of Financial Services Research 26, 245–261.

Akhigbe, Aigbe, Jeff Madura, and Ann Marie Whyte, 2004, Partial anticipation and the gains to bank merger targets, Journal of Financial Services Research 26, 55–71.

Alam, Ila M. Semenick, 2001, A nonparametric approach for assessing productivity dynamics of large U.S. banks, Journal of Money, Credit, and Banking 33, 121–139.

Alhadeff, David, 1974, Barriers to bank entry, Southern Economic Journal 40, 589–603.

D.V. Hoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-642-02821-2, © Springer-Verlag Berlin Heidelberg 2010

Alhadeff, David, 1954, Monopoly and Competition in Commercial Banking, Berkeley: University of California Press.

Ali, Ashiq, Sandy Klasa, and Eric Yeung, 2008, The impact of product market competition on a firm's disclosure policy and its information environment, Unpublished manuscript, University of Texas at Dallas, University of Arizona, and University of Georgia, August.

Allen, Franklin, Elena Carletti, and Robert Marquez, 2008, Credit market competition and capital regulation, Manuscript, University of Pennsylvania, Center for Financial Studies, and Arizona State University, January 13.

Allen, Franklin, and Douglas Gale, 2004, Competition and financial stability, Journal of Money, Credit, and Banking 36, 453–480.

Allen, Linda, and Aron Gottesman, 2006, The informational efficiency of the equity market as compared to the syndicated bank loan market, Journal of Financial Services Research 30, 5–42.

Allen, Linda, and Anoop Rai, 1996, Operational efficiency in banking: An international comparison, Journal of Banking and Finance 20, 655–672.

Allen, Linda, and Anthony Saunders, 1993, Forbearance and valuation of deposit insurance as a callable put, Journal of Banking and Finance 17, 629–643.

Allen, Linda, Anthony Saunders, and Gregory Udell, 1991, The pricing of retail deposits: Concentration and information, Journal of Financial Intermediation 1, 335–361.

Almazan, Andres, 2002, A model of competition in banking: Bank capital versus expertise, Journal of Financial Intermediation 11, 87–121.

Altunbas, Yener, Lynne Evans, and Philip Moyneux, 2001, Bank ownership and efficiency, Journal of Money, Credit, and Banking 33, 926–954.

Amel, Dean, and J. Nellie Liang, 1997, Determinants of entry and profits in local banking markets, Review of Industrial Organization 12, 59–78.

Amel, Dean, and Stephen Rhoades, 1989, Empirical evidence on the motives for bank mergers, Eastern Economic Journal 15, 17–27.

Amel, Dean, and Martha Starr-McCluer, 2002, Market definition in banking: Recent evidence, Antitrust Bulletin 47, 63–89.

American Banker, 2007, Ranking the banks: Bank and thrift holding companies with the most assets on December 31, 2006, May 3.

American Bar Association, 2007, Bank Mergers and Acquisitions Handbook, Section on Antitrust Law.

Anand, Bharat, and Alexander Galetovic, 2006, Relationships, competition, and the structure of investment banking markets, Journal of Industrial Economics 54, 151–199.

Anderlini, Luca, 1986, Central banks and moral hazard, Economic Theory Discussion Paper No. 103, Department of Applied Economics, University of Cambridge.

Angkinand, Apanard, 2009, Banking regulation and the output cost of banking crises, Journal of International Financial Markets, Institutions, and Money 19, 240–257.

Ariss, Rima Turk, 2008, Why do regulators encourage more market power in banking? Empirical evidence from developing countries, Manuscript, Lebanese American University.

Armstrong, Mark, 2006, Competition in two-sided markets, Rand Journal of Economics 37, 668–691.

Asaftei, Gabriel, 2008, The contribution of product mix versus efficiency and technical change in U.S. banking, Journal of Banking and Finance 32, 2336–2345.

Ashcraft, Adam, 2008a, Are bank holding companies a source of strength to their banking subsidiaries? Journal of Money, Credit, and Banking 40, 273–294.

Ashcraft, Adam, 2008b, Does the market discipline banks? New evidence from the regulatory capital mix, Journal of Financial Intermediation 17, 543–561.

Ashcraft, Adam. 2001, Do tougher bank capital requirements matter? New evidence from the eighties, Manuscript, Federal Reserve Bank of New York, July 23.

Ashcraft, Adam, and Hoyt Bleakley, 2006, On the market discipline of informationally opaque firms: Evidence from bank borrowers in the federal funds market, Federal Reserve Bank of New York Staff Report No. 257, August.

Ashton, John, and Khac Pham, 2007, Efficiency and price effects of horizontal bank mergers, University of East Anglia, ESRC Centre for Competition Policy Working Paper 07-9.

Avery, Robert, Raphael Bostic, Paul Calem, and Glenn Canner, 1999, Consolidation and bank branching patterns, Journal of Banking and Finance 23, 497–532.

Aydin, Halil, 2007, Interest rate pass-through in Turkey, Central Bank of the Republic of Turkey Working Paper No. 07/05.

Ayuso, Juan, Daniel Perez, and Jesus Saurina, 2004, Are capital buffers pro-cyclical? Evidence from Spanish panel data, Journal of Financial Intermediation 13, 249–264.

Bain, Joseph, 1956, Barriers to New Competition, Cambridge, MA: Harvard University Press.

Baltensperger, Ernst, 1980, Alternative approaches to the theory of the banking firm, Journal of Monetary Economics 6, 1–37.

Bank for International Settlements, 2006, International convergence of capital measurement and capital standards: A revised framework comprehensive version, June.

Barrios, Victor, and Jan Blanco, 2003, The effectiveness of bank capital adequacy regulation: A theoretical and empirical approach, Journal of Banking and Finance 27, 1935–1958.

Barro, Robert, and Anthony Santomero, 1972, Household money holdings and the demand deposit rate, Journal of Money, Credit, and Banking 4, 397–413.

Barros, P. Pita, 1999, Multimarket competition in banking, with an example from the Portuguese market, International Journal of Industrial Organization 17, 335–352.

Barth, James, 1991, The Great Savings and Loan Debacle, Washington, DC: American Enterprise Institute Press.

Barth, James, Gerard Caprio, Jr., and Ross Levine, 2008, Bank regulations are changing: For better or worse? World Bank Policy Research Paper 4646, June.

Barth, James, Gerard Caprio, Jr., and Ross Levine, 2006, Rethinking Bank Regulation, Cambridge: Cambridge University Press.

Barth, James, Gerard Caprio, Jr., and Ross Levine, 2004, Bank regulation and supervision: What works best? Journal of Financial Intermediation 13, 205–248.

Baxter, William, 1983, Bank interchanges of transactional paper: Legal and economic perspectives, Journal of Law and Economics 26, 541–588.

Beatty, Anne, and Anne Gron, 2001, Capital, portfolio, and growth: Bank behavior under risk-based capital requirements, Journal of Financial Services Research 20, 5–31.

Becher, David, and Terry Campbell II, 2005, Interstate banking deregulation and the changing nature of bank mergers, Journal of Financial Research 28, 1–20.

Beck, Thorsten, 2008, Bank competition and financial stability: Friends or foes? World Bank Policy Research Working Paper No. 4656. June.

Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine, 2006, Bank concentration, competition, and crises: First results, Journal of Banking and Finance 30, 1581–1603.

Becker, Gary, 1983, A theory of competition among pressure groups for political influence, Quarterly Journal of Economics 98, 371–400.

Becker, William, 1975, Determinants of the United States currency-demand deposit ratio, Journal of Finance 27, 57–74.

Benavie, Arthur, and Richard Froyen, 1982, Monetary policy in a model with a federal funds market: Fixed versus flexible deposit rates, Southern Economic Journal 48, 932–949.

Benston, George, 2000, Consumer protection as justification for regulating financial-services firms and products, Journal of Financial Services Research 17, 277–301.

Benston, George, 1965, Branch banking and economies of scale, Journal of Finance 20, 312–331.

Benston, George, Gerald Hanweck, and David Humphrey, 1982, Scale economies in banking: A restructuring and reassessment, Journal of Money, Credit, and Banking 14, 435–456.

Benston, George, William Hunter, and Larry Wall, 1995, Motivations for bank mergers and acquisitions: Enhancing the deposit insurance put option versus earnings diversification, Journal of Money, Credit, and Banking 27, 777–788.

Berg, Sigbjørn Alte, and Moshe Kim, 1998, Banks as multioutput oligopolies: An empirical evaluation of the retail and corporate banking markets, Journal of Money, Credit, and Banking 30, 135–153.

Berg, Sigbjørn Alte, and Moshe Kim, 1994, Oligopolistic interdependence and the structure of production in banking: An empirical investigation, Journal of Money, Credit, and Banking 26, 309–322.

Berger, Allen, 2006, Potential competitive effects of Basel II on banks in SME credit markets in the United States, Journal of Financial Services Research 29, 5–36.

Berger, Allen, 2003, The economic effects of technological progress: Evidence from the banking industry, Journal of Money, Credit, and Banking 35, 141–176.

Berger, Allen, 1995, The profit-structure relationship in banking: Tests of market power and efficient-structure hypotheses, Journal of Money, Credit, and Banking 27, 404–431.

Berger, Allen, Seth Bonime, Daniel Covitz, and Diana Hancock, 2000, Why are bank profits so persistent? The roles of product market competition, informational opacity, and regional/macroeconomic shocks, Journal of Banking and Finance 24, 1203–1235.

Berger, Allen, and Sally Davies, 1998, The information content of bank examinations, Journal of Financial Services Research 14, 117–144.

Berger, Allen, Sally Davies, and Mark Flannery, 2000, Comparing market and supervisory assessments of bank performance: Who knows what when? Journal of Money, Credit, and Banking 32, 641–667.

Berger, Allen, Rebecca Demsetz, and Philip Strahan, 1999, The consolidation of the financial services industry: Causes, consequences, and implications for the future, Journal of Banking and Finance 23, 135–194.

Berger, Allen, and Robert DeYoung, 2006, Technological progress and the geographic expansion of the banking industry, Journal of Money, Credit, and Banking 38, 1483–1513.

Berger, Allen, and Robert DeYoung, 2001, The effects of geographic expansion on bank efficiency, Journal of Financial Services Research 19, 163–184.

Berger, Allen, and Astrid Dick, 2007, Entry into banking markets and early-mover advantage, Journal of Money, Credit, and Banking 39, 775–807.

Berger, Allen, Astrid Dick, Lawrence Goldberg, and Lawrence White, 2007, Competition from large, multimarket firms and the performance of small, single-market firms: Evidence from the banking industry, Journal of Money, Credit, and Banking 39, 331–369.

Berger, Allen, and Timothy Hannan, 1998, The efficiency cost of market power in the banking industry: A test of the "quiet life" and related hypotheses, Review of Economics and Statistics 80, 454–465.

Berger, Allen, and Timothy Hannan, 1989, The price-concentration relationship in banking, Review of Economics and Statistics 71, 291–299.

Berger, Allen, Gerald Hanweck, and David Humphrey, 1987, Competitive viability in banking: Scale, scope, and product mix economies, Journal of Monetary Economics 20, 501–520.

Berger, Allen, Richard Herring, and Giorgio Szego, 1995, The role of capital in financial institutions. Journal of Banking and Finance 19, 393–430.

Berger, Allen, and David Humphrey, 1997, Efficiency of financial institutions: International survey and directions for future research, European Journal of Operational Research 98, 175–212.

Berger, Allen, and David Humphrey, 1991, The dominance of inefficiencies over scale and product mix economies in banking, Journal of Monetary Economics 28, 117–148.

Berger, Allen, David Humphrey, and Lawrence Pulley, 1996, Do consumers pay for one-stop banking? Evidence from an alternative revenue function, Journal of Banking and Finance 20, 1601–1621.

Berger, Allen, Leora Klapper, and Rima Turk-Ariss, 2008, Bank competition and financial stability, World Bank Policy Research Working Paper No. 4696, August.

Berger, Allen, and Loretta Mester, 2003, Explaining the dramatic change in performance of U.S. banks: Technological change, deregulation, and dynamic changes in competition, Journal of Financial Intermediation 12, 57–95.

Berger, Allen, and Loretta Mester, 1997, Inside the black box: What explains differences in the efficiencies of financial institutions? Journal of Banking and Finance 21, 895–947.

Berger, Allen, Richard Rosen, and Gregory Udell, 2007, Does market size structure affect competition? The case of small business lending, Journal of Banking and Finance 31, 11–33.

Berger, Allen, Anthony Saunders, Joseph Scalise, and Gergory Udell, 1998, The effects of bank mergers and acquisitions on small business lending, Journal of Financial Economics 50, 187–229.

Berger, Allen, and Gregory Udell, 2002, Small business credit availability and relationship lending: The importance of bank organizational structure, Economic Journal 112, F32–F53.

Berger, Allen, and Gregory Udell, 1995, Relationship lending and lines of credit in small firm finance, Journal of Business 68, 351–381.

Berger, Allen, and Gregory Udell, 1994, Did risk-based capital allocate bank credit and cause a 'credit crunch' in the United States? Journal of Money, Credit, and Banking 26, 585–628.

Berlin, Mitchell, and Loretta Mester, 1999, Deposits and relationship lending, Review of Financial Studies 12, 579–607.

Bernanke, Ben, 2001, Synergies between bank supervision and monetary policy: Comment, in Frederic Mishkin, ed., Prudential Supervision: What Works and What Doesn't, Chicago: University of Chicago Press, pp. 293–300.

Berry, Steven, 1994, Estimating discrete-choice models of product differentiation, Rand Journal of Economics 25, 242–262.

Berry, Steven, James Levinsohn, and Ariel Pakes, 1995, Automobile prices in market equilibrium, Econometrica 63, 841–890.

Berstein, Solange, and J. Rodrigo Fuentes, 2004, Is there lending rate stickiness in the Chilean banking industry? in Luis Antonio Ahumada and J. Rodrigo Fuentes, eds., Banking Market Structure and Monetary Policy, Santiago: Central Bank of Chile, pp. 183–210.

Besanko, David, and George Kanatas, 1996, The regulation of bank capital: Do capital standards promote bank safety? Journal of Financial Intermediation 5, 160–183.

Besanko, David, and Anjan Thakor, 1992, Banking deregulation: Allocational consequences of relaxing entry barriers, Journal of Banking and Finance 16, 909–932.

Bhattacharya, Sudipto, Arnoud Boot, and Anjan Thakor, 1998, The economics of bank regulation, Journal of Money, Credit, and Banking 30, 745–770.

Bikker, Jacob, and Katarina Haaf, 2002, Competition, concentration, and their relationship: An empirical analysis of the banking industry, Journal of Banking and Finance 26, 2191–2214.

Billett, Matthew, Mark Flannery, and Jon Garfinkel, 1995, The effect of lender identity on a borrowing firm's equity return, Journal of Finance 50, 699–718.

Billet, Matthew, Jon Garfinkel, and Edward O'Neal, 1998, The cost of market versus regulatory discipline in banking, Journal of Financial Economics 48, 333–358.

Birchler, Urs, and Andréa Maechler, 2001, Do depositors discipline Swiss banks? Swiss National Bank Working Paper No. 1.06.

Blackwell, David, and Drew Winters, 1997, Banking relationships and the effect of monitoring on loan pricing, Journal of Financial Research 20, 275–289.

Blair, Roger, and Arnold Heggestad, 1978, Bank portfolio regulation and the probability of bank failure, Journal of Money, Credit, and Banking 10, 88–93.

Bliss, Robert, 2008, Multiple regulators and insolvency regimes: Obstacles to efficient supervision and resolution, in David Mayes and Geoffrey Wood, eds., The Structure of Financial Regulation, London: Routledge, pp. 132–154.

Bliss, Robert, 2004, Market discipline: Players, processes, and purposes, in Claudio Borio, William Hunter, George Kaufman, and Kostas Tsatsaronis, eds., Market Discipline across Countries and Industries, Cambridge, MA: MIT Press, pp. 37–53.

Bliss, Robert, 2001, Market discipline and subordinated debt: A review of some salient issues, Federal Reserve Bank of Chicago Economic Perspectives, First Quarter 25, 24–45.

Bliss, Robert, and Mark Flannery, 2000, Market discipline in the governance of U.S. bank holding companies: Monitoring versus influencing, Federal Reserve Bank of Chicago and University of Florida, Research Department Working Paper No. 2000-03.

Bliss, Robert, and George Kaufman, 2003, Bank procyclicality, credit crunches, and asymmetric monetary policy effects: A unifying model, Journal of Applied Finance 13 (Fall/Winter), 23–31.

Blum, Jürg, 2008, Why 'Basel II' may need a leverage ratio restriction, Journal of Banking and Finance 32, 1699–1707.

Blum, Jürg, 2002, Subordinated debt, market discipline, and banks' risk taking, Journal of Banking and Finance 26, 1427–1441.

Blum, Jürg, 1999, Do capital adequacy requirements reduce risks in banking? Journal of Banking and Finance 23, 755–771.

Bolt, Wilko, and Alexander Tieman, 2006, Social welfare and cost recovery in two-sided markets, Review of Network Economics 5, 103–117.

Bolton, Patrick, Xavier Freixas, and Joel Shapiro, 2004, Conflicts of interest, information provision, and competition in banking, National Bureau of Economic Research Working Paper No. 10571, June.

Bond, Philip, Itay Goldstein, and Edward Simpson Prescott, 2008, Market-based corrective action, Manuscript, University of Pennsylvania and Federal Reserve Bank of Richmond, July.

Boone, Jan, 2001, Intensity of competition and the incentive to innovate, International Journal of Industrial Organization 19, 705–726.

Boot, Arnoud, 2000, Relationship banking: What do we know? Journal of Financial Intermediation 9, 7–25.

Boot, Arnoud, and Matej Marinč, 2008, The evolving landscape of banking, Industrial and Corporate Change 17, 1173–1203.

Boot, Arnoud, and Matej Marinč, 2006, Competition and entry in banking: Implications for stability and capital regulation, Tinbergen Institute Discussion Paper TI 2006-015/2.

Boot, Arnoud, and Anjolein Schmeits, 2000, Market discipline and incentive problems in conglomerate firms with applications to banking, Journal of Financial Intermediation 9, 240–273.

Boot, Arnoud, and Anjan Thakor, 2001, The many faces of information disclosure, Review of Financial Studies 14, 1021–1057.

Boot, Arnoud, and Anjan Thakor, 2000, Can relationship banking survive competition? Journal of Finance 55, 679–713.

Boot, Arnoud, and Anjan Thakor, 1993, Self-interested bank regulation, American Economic Review Papers and Proceedings 83, 206–212.

Bordo, Michael, 2003, Market discipline and financial crisis policy: A historical perspective, in George Kaufman, ed., Market Discipline in Banking: Theory and Evidence, Amsterdam: Elsevier, pp. 157–182.

Borio, Claudio, 2003, Towards a macroprudential framework for financial supervision and regulation, CESifo Economics Studies 49, 181–215.

Bossone, Biagio, and Jong-Kun Lee, 2004, In finance, size matters: The "systemic scale economies" hypothesis, IMF Staff Papers 51, 19–46.

Bouvatier, Vincent, and Laetitia Lepetit, 2008, Banks' procyclical behavior: Does provisioning matter? Journal of International Financial Markets, Institutions, and Money 18, 513–526.

Boyd, John, and Gianni De Nicolo, 2005, The theory of bank risk taking and competition revisited, Journal of Finance 60, 1329–1343.

Boyd, John, Gianni De Nicolo, and Abu Jalal, 2006, Bank risk-taking and competition revisited: New theory and new evidence, IMF Working Paper WP/06/297, December.

Boyd, John, and Mark Gertler, 1993, U.S. commercial banks: Trends, cycles, and policy, in Olivier Blanchard and Stanley Fischer, eds., NBER Macroeconomics Annual, Cambridge, MA: MIT Press, pp. 319–368.

Boyd, John, and Graham, Stanley, 1996, Consolidation in U.S. banking: Implications for efficiency and risk, Federal Reserve Bank of Minneapolis Working Paper 572, December.

Bradley, Michael D., and Dennis Jansen, 1986, On deposit market deregulation and interest rates, Southern Economic Journal 53, 478–489.

Bradley, Michael, Carol Wambeke, and David Whidbee, 1991, Risk weights, risk-based capital, and deposit insurance, Journal of Banking and Finance 15, 875–893.

Brevoort, Kenneth, and Timothy Hannan, 2006, Commercial lending and distance: Evidence from community reinvestment act data, Journal of Money, Credit, and Banking 38, 1991–2012.

Brewer, Elijah III, and William Jackson III, 2006, A note on the "risk-adjusted" price-concentration relationship in banking, Journal of Banking and Finance 30, 1041–1054.

Brinkmann, Emile and Paul Horvitz, 1995, Risk-based capital standards and the credit crunch, Journal of Money, Credit, and Banking 27, 848–863.

Brissimis, Sophocles, and Manthos Delis, 2009, Bank-level estimates of market power, Bank of Greece and University of Ioannina, Working Papers 93, January.

Broecker, Thorsten, 1990, Creditworthiness tests and interbank competition, Econometrica 58, 429–452.

Budzinski, Oliver, 2008, A note on competing merger simulation models in antitrust cases: Can the best be identified? Marburg Paper on Economics 2008-01, Marburg Center for Institutional Economics, Philipps-University of Marburg, March 27.

Budzinski, Oliver, and Isabel Ruhmer, 2008, Merger simulation in competition policy: A survey, MAGKS Joint Discussion Paper No. 07-2008.

Burgstaller, Johann, 2005, Interest rate pass-through estimates from vector autoregressive models, Unpublished Manuscript, Johannes Kepler University.

Cabral, Luis, 2005, Market power and efficiency in card payment systems: A comment, Review of Network Economics 5, 15–25.

Calem, Paul, and Gerald Carlino, 1991, The concentration/conduct relationship in bank deposit markets, Review of Economics and Statistics 73, 268–276.

Calem, Paul, and Leonard Nakamura, 1998, Branch banking and the geography of bank pricing, Review of Economics and Statistics 80, 600–610.

Calem, Paul, and Rafael Rob, 1999, The impact of capital-based regulation on bank risk-taking, Journal of Financial Intermediation 8, 317–352.

Calomiris, Charles, 2004, Reestablishing market discipline as a part of bank regulation, in Claudio Borio, William Hunter, George Kaufman, and Kostas Tsatsaronis, eds., Market Discipline across Countries and Industries, Cambridge, MA: MIT Press, pp. 407–421.

Calomiris, Charles, 1999, Building an incentive-compatible safety net, Journal of Banking and Finance 23, 1499–1519.

Calomiris, Charles, and Charles Kahn, 1991, The role of demandable debt in structuring optimal banking arrangements, American Economic Review 81, 497–513.

Calomiris, Charles, and Andrew Powell, 2000, Can emerging market bank regulators establish credible discipline? The case of Argentina, 1992–1999, NBER Working Paper No. 7715.

Caminal, Ramon, and Carmen Matutes, 2002, Market power and banking failures, International Journal of Industrial Organization 20, 1341–1361.

Campa, José Manuel, and Ignacio Hernando, 2006, M&As performance in the European financial industry, Journal of Banking and Finance, 3367–3392.

Campbell, Tim, Yuk-Shee Chan, and Anthony Marino, 1992, An incentive-based theory of bank regulation, Journal of Financial Intermediation 2, 255–276.

Cannata, Francesco, and Mario Quagliariello, 2005, The value of market information in banking supervision: Evidence from Italy, Journal of Financial Services Research 27, 139–162.

Carbó-Valverde, Santiago, David Humphrey, Joaquin Maudos, and Philip Molyneux, 2009, Cross-country comparisons of competition and pricing power in European banking, Journal of International Money and Finance 28, 115–134.

Carbó-Valverde, Santiago, Timothy Hannan, and Francisco Rodriguez-Fernandez, 2009, Exploiting old customers and attracting new ones: The case of bank deposit pricing, Presented at the International Industrial Organization Society Conference, Boston, MA, April.

Carey, Mark. 2002, A guide to choosing absolute bank capital requirements, Journal of Banking and Finance 26, 929–951.

Carletti, Elena, 2008, Competition and regulation in banking, in Anjan Thakor and Arnoud Boot, eds., Handbook of Financial Intermediation and Banking, Amsterdam: Elsevier, pp. 449–482.

Carletti, Elena, 2004, The structure of bank relationships, endogenous monitoring, and loan rates, Journal of Financial Intermediation 13, 58–86.

Carletti, Elena, Vittoria Cerasi, and Sonja Daltung, 2007, Multiple-bank lending: Diversification and free-riding in monitoring, Journal of Financial Intermediation 16, 425–451.

Carletti, Elena, Philipp Hartmann, and Giancarlo Spagnolo, 2007, Bank mergers, competition, and liquidity, Journal of Money, Credit, and Banking 39, 1067–1105.

Carletti, Elena, Philipp Hartmann, and Giancarlo Spagnolo, 2002, Implications of the bank merger wave for competition and stability, Third Joint Central Bank Research Conference on Risk Measurement and Systemic Risk, Bank for International Settlements, March 7–8.

Carlson, Mark, and Kris James Mitchener, 2006, Branch banking, bank competition, and financial stability, Journal of Money, Credit, and Banking 38, 1293–1328.

Carow, Kenneth, Edward Kane, and Rajesh Narayanan, 2006, How have borrowers fared in banking megamergers? Journal of Money, Credit, and Banking 38, 821–836.

Carter, David, and James McNulty, 2005, Deregulation, technological change, and the business-lending performance of large and small banks, Journal of Banking and Finance 29, 1113–1130.

Casolaro, Luca, Dario Focarelli, and Alberto Franco Pozzolo, 2003, The pricing effect of certification on bank loans: Evidence from the syndicated credit market, Economics and Statistics Discussion Paper No. 10/03, University of Molise.

Catarieneu-Rabell, Eva, Patricia Jackson, and Dimitrios Tsomocos, 2005, Procyclicality and the new Basel accord: Banks' choice of loan rate system, Economic Theory 26, 537–557.

Cerasi, Vittoria, Barbara Chizzolini, and Marc Ivaldi, 2002, Branching and competition in the European banking industry, Applied Economics 34, 2213–2225.

Chakravorti, Sujit, 2003, Theory of credit card networks: A survey of the literature, Review of Network Economics 2, 50–68.

Chakravorti, Sujit, and Roberto Roson, 2006, Platform competition in two-sided markets: The case of payment networks, Review of Network Economics 5, 118–140.

Chakravorti, Sujit, and Ted To, 2007, A theory of credit cards, International Journal of Industrial Organization, 583–595.

Chamberlin, Edward, 1962, The Theory of Monopolistic Competition, Cambridge, MA: Harvard University Press.

Chen, Andrew, Nengjiu Ju, Sumon Mazumdar, and Avinash Verma, 2006, Correlated default risks and bank regulations, Journal of Money, Credit, and Banking 38, 375–398.

Chen, Yehning, and Iftekhar Hasan, 2006, The transparency of the banking system and the efficiency of information-based bank runs, Journal of Financial Intermediation 15, 307–331.

Cheng, David, Benton Gup, and Larry Wall, 1989, Financial determinants of bank takeovers, Journal of Money, Credit, and Banking 21, 524–536.

Chiappori, Pierre-André, David Perez-Castrillo, and Thierry Verdier, 1995, Spatial competition in the banking system: Localization, cross subsidies, and the regulation of deposit rates, European Economic Review 39, 889–918.

Chiuri, Maria Concetta, Giovanni Ferri, and Giovanni Majnoni. 2002. The macroeconomic impact of bank capital requirements in emerging economies: Past evidence to assess the future. Journal of Banking and Finance 26, 881–904.

Čihák, Martin, and Richard Podpiera, 2008, Integrated financial supervision: Which model? North American Journal of Economics and Finance 19, 135–152.

Claessens, Stijn, and Luc Laeven, 2004, What drives bank competition? Some international evidence, Journal of Money, Credit, and Banking 36, 563–583.

Clark, Jeffrey, 1984, Estimation of economies of scale in banking using a generalized functional form, Journal of Money, Credit, and Banking 16, 53–68.

Clark, Jeffrey, and Thomas Siems, 2002, X-efficiency in banking: Looking beyond the balance sheet, Journal of Money, Credit, and Banking 34, 987–1013.

Clark, Timothy, Astrid Dick, Beverly Hirtle, Kevin Stiroh, and Robard Williams, 2007, The role of retail banking in the U.S. banking industry: Risk, return, and industry structure, Federal Reserve Bank of New York Economic Policy Review 14, 39–56.

Coccorese, Paolo, 2009, Market power in local banking monopolies, Journal of Banking and Finance 33, 1196–1210.

Coccorese, Paolo, 2005, Competition in markets with dominant firms: A note on the evidence from the Italian banking industry, Journal of Banking and Finance 29, 1083–1093.

Cole, Rebel, and Jeffrey Gunther, 1998, Predicting bank failures: A comparison of on- and off-site monitoring systems, Journal of Financial Services Research 13, 103–117.

Coleman, Anthony, Neil Esho, and Ian Sharpe, 2006, Does bank monitoring influence loan contract terms, Journal of Financial Services Research 30, 177–198.

Colwell, R.J., and E. Philip Davis, 1992, Output and productivity in banking, Scandinavian Journal of Economics 94, S111–S129.

Cooper, Russell, and Thomas Ross, 2002, Bank runs, deposit insurance, and capital requirements, International Economic Review 43, 55–71.

Cordell, Lawrence, and Kathleen Kuester King, 1995, A market evaluations of the risk-based capital standards for the U.S. financial system, Journal of Banking and Finance 19, 531–562.

Cordella, Tito, and Eduardo Yeyati, 2002, Financial opening, deposit insurance, and risk in a model of banking competition, European Economic Review 46, 471–485.

Cornett, Marcia Millon, Jamie John McNutt, and Hassan Tehranian, 2006, Performance changes around bank mergers: Revenue enhancements versus cost reductions, Journal of Money, Credit, and Banking 38, 1013–1050.

Corvoisier, Sandrine, and Reint Gropp, 2002, Bank concentration and retail interest rates, Journal of Banking and Finance 26, 2155–2189.

Cosimano, Thomas, 1988, The banking industry under uncertain monetary policy, Journal of Banking and Finance 12, 117–139.

Cosimano, Thomas, 1987, The federal funds market under bank deregulation, Journal of Money, Credit, and Banking 19, 326–339.

Cosimano, Thomas, and Van Huyck, John, 1989, Dynamic monetary control and interest rate stabilization, Journal of Monetary Economics 23, 53–63.

Cottarelli, Carlo, and Angeliki Kourelis, 1994, Financial structure, bank lending rates, and the transmission mechanism of monetary policy, IMF Staff Papers 41, 587–623.

Covitz, Daniel, Diana Hancock, and Myron Kwast, 2004, Market discipline in banking reconsidered: The roles of funding managers' decisions and deposit insurance reform, Board of Governors of the Federal Reserve System Financial and Economics Discussion Series No. 2004-53.

Craig, Steven, and Pauline Hardee, 2007, The impact of bank consolidation on small business credit availability, Journal of Banking and Finance 31, 1237–1263.

Cukierman, Alex, 1992, Central bank strategy, credibility, and independence: Theory and evidence, Cambridge, MA: MIT Press.

Curry, Timothy, Gary Fissel, and Carlos Ramirez, 2008, The impact of bank supervision on loan growth, North American Journal of Economics and Finance 19, 113–134.

Cybo-Ottone, Alberto, and Maurizio Murgia, 2000, Mergers and shareholder wealth in European banking, Journal of Banking and Finance 24, 831–859.

Dahiya, Sandeep, Manju Puri, and Anthony Saunders, 2003, Bank borrowers and loan sales: New evidence on the uniqueness of bank loans, Journal of Business 76, 563–582.

Dasgupta, Partha, and Joseph Stiglitz, 1981, Entry, innovation, exit: Towards a dynamic theory of industrial market structure, European Economic Review 15, 137–158.

Datta, Sudip, Mai Iskandar-Datta, and Ajay Patel, 1999, Bank monitoring and the pricing of corporate debt, Journal of Financial Economics 51, 435–449.

Davenport, Andrew, and Kathleen McDill, 2006, The depositor behind the discipline: A micro-level case study of Hamilton Bank, Journal of Financial Services Research 30, 93–109.

De Bondt, Gabe, 2005, Interest rate pass-through: Empirical results for the euro area," German Economic Review 6, 37–78.

De Bondt, Gabe, Benoit Mojon, and Natacha Valla, 2005, Term structure and the sluggishness of retail bank interest rates in euro area countries, European Central Bank Working Paper No. 518, September.

Decamps, Jean-Paul, Jean-Charles Rochet, and Benoit Roger, 2004, The three pillars of Basel II: Optimizing the mix, Journal of Financial Intermediation 13, 132–155.

Degryse, Hans, Moshe Kim, and Steven Ongena, 2009, Microeconometrics of Banking: Methods, Applications, and Results, Oxford: Oxford University Press.

Degryse, Hans, and Steven Ongena, 2008, Competition and regulation in the banking sector: A review of the empirical evidence on the sources of bank rents, in Anjan Thakor and Arnoud Boot, eds., Handbook of Financial Intermediation and Banking, Amsterdam: Elsevier, pp. 483–554.

Degryse, Hans, and Steven Ongena, 2007, The impact of competition on bank orientation, Journal of Financial Intermediation 16, 399–424.

Degryse, Hans, and Steven Ongena, 2005, Distance, lending relationships, and competition, Journal of Finance 60, 231–266.

De Guevara, Juan Fernández, Joaquín Maudos, and Francisco Pérez, 2005, Market power in European banking sectors, Journal of Financial Services Research 27, 109–137.

De Jonghe, Olivier, and Rudi Vander Vennet, 2008, Competition versus efficiency: What drives franchise values in European banking? Journal of Banking and Finance 32, 1820–1835.

De la Torre, Augusto, Maria Soledad Martínez Pería, and Sergio Schmukler, 2008, Bank involvement with SMEs: Beyond relationship lending, World Bank Policy Research Working Paper No. 4649, June.

Delis, Manthos, Philip Molyneux, and Fotios Pasiouras, 2008, Regulations and productivity growth in banking, Manuscript, Athens University of Economics and Business, Bangor University, and University of Bath.

Dell'Ariccia, Giovanni, 2001, Asymmetric information and the structure of the banking industry, European Economic Review 45, 1957–1980.

Dell'Ariccia, Giovanni, Ezra Friedman, and Robert Marquez, 1999, Adverse selection as a barrier to entry in the banking industry, Rand Journal of Economics 30, 515–534.

Dell'Ariccia, Giovanni, and Robert Marquez, 2006, Competition among regulators and credit market integration, Journal of Financial Economics 79, 401–430.

DeLong, Gayle, 2003, Does long-term performance of mergers match market expectations? Evidence from the U.S. banking industry, Financial Management 32 (Summer), 5–25.

DeLong, Gayle, 2001, Stockholder gains from focusing versus diversifying bank mergers, Journal of Financial Economics 59, 221–252.

Demirgüç-Kunt, Asli, Enrica Detragiache, and Thierry Tressel, 2008, Banking on the principles: Compliance with Basel core principles and bank soundness, Journal of Financial Intermediation 17, 511–542.

Demirgüç-Kunt, Asli, and Harry Huizinga, 2009, Bank activity and funding strategies: The impact on risk and return, Tilburg University, European Banking Center Discussion Paper 2009-09, January.

Demirgüç-Kunt, Asli, Edward Kane, and Luc Laeven, eds., 2008, Deposit Insurance Around the World: Issues of Design and Implementation, Cambridge, MA: MIT Press.

Demirgüç-Kunt, Asli, Luc Laeven, and Ross Levine, 2004, Regulations, market structure, institutions, and the cost of financial intermediation, Journal of Money, Credit, and Banking 36, 593–622.

Demsetz, Harold, 1973, Industry structure, market rivalry, and public policy, Journal of Law and Economics 16, 1–9.

Deng, Saiying (Esther), Elyas Elyasiani, and Connie Mao, 2007, Diversification and the cost of debt at bank holding companies, Journal of Banking and Finance 31, 2453–2473.
De Nicolò, Gianni, 2001, Size, charter value, and risk in banking: An international perspective, Manuscript, International Monetary Fund, April.
De Nicolò, Gianni, and Elena Loukoianova, 2007, Bank ownership, market structure, and risk, IMF Working Paper No. WP/07/215.
de Pinho, Paolo, 2000, The impact of deregulation on price and non-price competition in the Portuguese deposits market, Journal of Banking and Finance 24, 1515–1533.
Devaney, Michael, and William Weber, 2002, Small-business lending and profit efficiency in commercial banking, Journal of Financial Services Research 22, 225–246.
De Vries, Casper George, 2005, The simple economics of bank fragility, Journal of Banking and Finance 29, 803–825.
Dewatripont, Mathias, and Jean Tirole, 1993, The Prudential Regulation of Banks, Cambridge, MA: MIT Press.
DeYoung, Robert, 2008, Safety, soundness, and the evolution of the U.S. banking industry, in Anjan Thakor and Arnoud Boot, eds., Handbook of Financial Intermediation and Banking, Amsterdam: Elsevier, pp. 347–373.
DeYoung, Robert, 2003, De novo bank exit, Journal of Money, Credit, and Banking 35, 711–728.
DeYoung, Robert, 1998, Management quality and X-inefficiency in national banks, Journal of Financial Services Research 13, 5–22.
DeYoung, Robert, William Curt Hunter, and Gregory Udell, 2004, The past, present, and probable future for community banks, Journal of Financial Services Research 25, 85–133.
DeYoung, Robert, and Evren Örs, 2004, Advertising and pricing at multiple-output firms: Evidence from U.S. thrift institutions, Federal Reserve Bank of Chicago Working Paper WP 2004-25.
DeYoung, Robert, and Karin Roland, 2001, Product mix and earnings volatility at commercial banks: Evidence from a degree of total leverage model, Journal of Financial Intermediation 10, 54–84.
DeYoung, Robert, and Chiwon Yom, 2008, On the independence of assets and liabilities: Evidence from U.S. commercial banks, 1990–2005, Journal of Financial Stability 4, 275–303.
Diamond, Douglas, 1996, Financial intermediation as delegated monitoring: A simple example. Federal Reserve Bank of Richmond Economic Quarterly 82/3 (Summer), 51–66.
Diamond, Douglas, 1991, Monitoring and regulation: The choice between bank loans and directly placed debt, Journal of Political Economy 99, 689–721.
Diamond, Douglas, 1984, Financial intermediation and delegated monitoring, Review of Economic Studies 51, 393–414.
Diamond, Douglas, and Phillip Dybvig, 1983, Bank runs, deposit insurance, and liquidity, Journal of Political Economy 91, 401–419.
Diamond, Douglas, and Ragharam Rajan, 2006, Money in a theory of banking, American Economic Review 96, 30–53.
Diamond, Douglas, and Ragharam Rajan, 2001, Liquidity risk, liquidity creation, and financial fragility: A theory of banking, Journal of Political Economy 109, 287–327.
Diamond, Douglas, and Raghuram Rajan, 2000, A theory of bank capital, Journal of Finance 55, 2431–2465.
Dick, Astrid, 2008, Demand estimation and consumer welfare in the banking industry, Journal of Banking and Finance 32, 1661–1676.
Dick, Astrid, 2007, Market size, service quality, and competition in banking, Journal of Money, Credit, and Banking 39, 49–81.
Dick, Astrid, 2006, Nationwide branching and its impact on market structure, quality, and bank performance, Journal of Business 79, 567–592.
Dietrich, J. Kimball, and Christopher James, 1983, Regulation and the determination of bank capital changes: A note, Journal of Finance 38, 1651–1658.

Dinç, L. Serdar, 2000, Bank reputation, bank commitment, and the effects of competition in credit markets, Review of Financial Studies 13, 781–812.

Di Noia, Carmine, and Giorgio Di Giorgio, 1999, Should banking supervision and monetary policy tasks be given to different agencies? International Finance 2, 361–378.

Di Patti, Emilia Bonaccorsi, and Giogio Gobbi, 2007, Winners or losers? The effects of banking consolidation on corporate borrowers, Journal of Finance 62, 669–695.

Distinguin, Isabelle, Philippe Rous, and Amine Tarazi, 2006, Market discipline and the use of stock market data to predict bank financial distress, Journal of Financial Services Research 30, 151–176.

Dowd, Kevin, 2000, Bank capital adequacy versus deposit insurance, Journal of Financial Services Research 17, 7–15.

Dowd, Kevin, 1994, Competitive banking, bankers' clubs, and bank regulation, Journal of Money, Credit, and Banking 26, 289–308.

Dowd, Kevin, 1992, Models of banking instability: A partial review of the literature, Journal of Economic Surveys 6, 107–132.

Dreyfus, Jean-Francois, Anthony Saunders, and Linda Allen, 1994, Deposit insurance and regulatory forbearance: Are caps on insured deposits optimal? Journal of Money, Credit, and Banking 26, 412–438.

Economides, Nicholas, 2008, Competition policy issues in the consumer payments industry, Law and Economics Research Working Paper No. 08-56, New York University School of Law, November.

Edelstein, Paul, and Donald Morgan, 2006, Local or state? Evidence on bank market size using branch prices, Federal Reserve Bank of New York Economic Policy Review 13, 15–25.

Edwards, Franklin, 1976, More on advertising and competition in banking, Antitrust Bulletin 21, 85–89.

Edwards, Franklin, 1973, Advertising and competition in banking, Antitrust Bulletin 18, 23–31.

Égert, Balázs, Jesús Crespo-Cuaresma, and Thomas Reininger, 2007, Interest rate pass-through in central and eastern Europe: Reborn from ashes merely to pass away? Journal of Policy Modeling 19, 207–225.

Elliehausen, Gregory, 1998, The cost of bank regulation: A review of the evidence, Staff Study 171, Board of Governors of the Federal Reserve System, April.

Elsas, Ralf, 2005, Empirical determinants of relationship lending, Journal of Financial Intermediation 14, 32–57.

Elyasiani, Elyas, Kenneth Kopecky, and David VanHoose, 1995, Costs of adjustment, portfolio separation, and the dynamic behavior of bank loans and deposits, Journal of Money, Credit, and Banking 27, 955–974.

Emch, Eric, and T. Scott Thompson, 2006, Market definition and market power in payment card networks, Review of Network Economics 5, 45–60.

Emmons, William, R. Alton Gilbert, and Timothy Yeager, 2004, Reducing the risk at small community banks: Is it size or geographic diversification that matters? Journal of Financial Services Research, 259–281.

Ennis, Huberto, and H. Sam Malek, 2005, Bank risk of failure and the too-big-to-fail policy, Federal Reserve Bank of Richmond Economic Quarterly, 91 (Spring), 21–44.

Erel, Isil, 2007, The effect of bank mergers on loan prices: Evidence from the United States, Dice Center Working Paper WP 2006-19, December.

Ergungor, O. Emre, 2005, The profitability of bank-borrower relationships, Journal of Financial Intermediation 14, 485–512.

Ernst and Young, 2006, Basel II: The business impact of Basel II sinking in as Canadian banks struggle with implementation, April 12. (http://www.ey.com/global/content.nsf/Canada/Media_-_2006_-_Basel_II)

Espinoza-Vega, Marco, and Allesandro Rebucci, 2003, Retail bank interest rate pass-through: Is Chile atypical? IMF Working Paper WP/03/112.

Estrella, Arturo, 2004a, Bank capital and risk: Is voluntary disclosure enough? Journal of Financial Services Research 26, 145–160.

Estrella, Arturo, 2004b, The cyclical behavior of optimal bank capital, Journal of Banking and Finance 28, 1469–1498.

Esty, Benjamin, Bhanu Narasimhan, and Peter Tufano, 1999, Interest-rate exposure and bank mergers, Journal of Banking and Finance 23, 255–285.

Evanoff, Douglas, and Evren Örs, 2009, The competitive dynamics of geographic deregulation in banking: Implications for productive efficiency, Journal of Money, Credit, and Banking, 40, 847–928.

Evanoff, Douglas, and Larry Wall, 2001, Sub-debt yield spreads as bank risk measures, Journal of Financial Services Research 20, 121–145.

Evanoff, Douglas, and Larry Wall, 2000, Subordinated debt as bank capital: A proposal for regulatory reform, Federal Reserve Bank of Chicago Economic Perspectives, 40–53.

Evans, David, and Richard Schmalensee, 1999, Paying with Plastic: The Digital Revolution in Buying and Borrowing, Cambridge, MA: MIT Press.

Felici, Roberto, and Marcello Pagnini, 2008, Distance, bank heterogeneity, and entry in local banking markets, Journal of Industrial Economics 56, 500–534.

Fields, L. Paige, Donald Fraser, Tammy Berry, and Steven Byers, 2006, Do bank loan relationships still matter? Journal of Money, Credit, and Banking 38, 1195–1209.

Flannery, Mark, 2001, The faces of "market discipline," Journal of Financial Services Research 20, 107–119.

Flannery, Mark, 1998, Using market information in prudential supervision: A review of the U.S. empirical evidence, Journal of Money, Credit, and Banking 30, 273–305.

Flannery, Mark, 1989, Capital regulation and insured banks' choice of individual loan default risks, Journal of Monetary Economics 24, 235–258.

Flannery, Mark, 1982, Retail bank deposits as quasi-fixed factors of production, American Economic Review 72, 527–536.

Flannery, Mark, Simon Kwan, and M. Nimalendran, 2004, Market evidence on the opaqueness of banking firms' assets, Journal of Financial Economics 71, 419–460.

Flannery, Mark, and Kasturi Rangan, 2004, What caused the bank capital build-up of the 1990s? FDIC Center for Financial Research Working Paper No. 2004-03, August.

Flannery, Mark, and Sorin Sorescu, 1996, Evidence of bank market discipline in subordinated debenture yields: 1983–1991, Journal of Finance 51, 1347–1377.

Focarelli, Dario, and Fabio Panetta, 2003, Are mergers beneficial to consumers? Evidence from the market for bank deposits, American Economic Review 93, 1152–1171.

Focarelli, Dario, Fabio Panetta, and Carmelo Salleo, 2002, Why do banks merge? Journal of Money, Credit, and Banking 34, 1047–1066.

Franck, Raphaël, and Miriam Krausz, 2008, Why separate monetary policy from banking supervision? Journal of Comparative Economics 36, 388–411.

Freeman, Scott, 1988, Banking as the provision of liquidity, Journal of Business 61, 45–64.

Freixas, Xavier, and Bruno Parigi, 2007, Banking regulation and prompt corrective action, Manuscript, Universitat Pompeu Fabra and Univeristy of Padova, January 11.

Freixas, Xavier, and Jean-Charles Rochet, 2008, Microeconomics of Banking, 2nd Edition, Cambridge, MA: MIT Press.

Fries, Steven, Pierre Mella-Barral, and William Perraudin, 1997, Optimal bank reorganization and the fair pricing of deposit guarantees, Journal of Banking and Finance 21, 441–468.

Furfine, Craig, 2001, Bank portfolio allocation: The impact of capital requirements, regulatory monitoring, and economic conditions, Journal of Financial Services Research 20, 33–56.

Furlong, Frederick, 1992, Capital regulation and bank lending, Federal Reserve Bank of San Francisco Economic Review, 23–33.

Furlong, Frederick and Michael Keeley, 1989, Capital regulation and bank risk-taking: A note. Journal of Banking and Finance 13, 883–891.

Gande, Amar, Manu Puri, Anthony Saunders, and Ingo Walter, 1997, Bank underwriting of debt securities: Modern evidence, Review of Financial Studies 10, 1175–1202.

Garmaise, Mark, and Tobias Moskowitz, 2006, Bank mergers and crime: The real and social effects of credit market competition, Journal of Finance 61, 495–538.

Gennotte, Gerard and David Pyle, 1991, Capital controls and bank risk, Journal of Banking and Finance 15, 805–824.

Gilbert, R. Alton, 1990, Market discipline of bank risk: Theory and evidence, Federal Reserve Bank of St. Louis Review, January–February, 72, 3–18.

Gilbert, R. Alton, 1984, Bank market structure and competition: A survey, Journal of Money, Credit, and Banking 16, 617–644.

Gilbert, R. Alton, and Adam Zaretsky, 2003, Banking antitrust: Are the assumptions still valid? Federal Reserve Bank of St. Louis Review, November/December, 85, 29–52.

Gilligan, Thomas, William Marshall, and Michael Smirlock, 1984, Scale and scope economies in the multiproduct banking firm, Journal of Monetary Economics 13, 393–405.

Gilligan, Thomas, and Michael Smirlock, 1984, An empirical study of joint production and scale economies in commercial banking, Journal of Banking and Finance 8, 67–77.

Gjerde, Øystein, and Kristian Semmen, 1995, Risk-based capital requirements and bank portfolio risk, Journal of Banking and Finance 19, 1159–1173.

Glassman, Cynthia, and Stephen Rhoades, 1980, Owner vs. manager effects on bank performance, Review of Economics and Statistics 62, 263–270.

Gobbi, Giorgio, and Francesca Lotti, 2004, Entry decisions and adverse selection: An empirical analysis of local credit markets, Journal of Financial Services Research 26, 225–244.

Goddard, John, Phil Molyneux, and John Wilson, 2004, Dynamics of growth and profitability in banking, Journal of Money, Credit, and Banking 36, 1069–1090.

Goldstein, Itay, and Ady Pauzner, 2005, Demand deposit contracts and the probability of bank runs, Journal of Finance 60, 1293–1328.

Goodhart, Charles, 2000, The organizational structure of banking supervision, Financial Stability Institute Occasional Paper No. 1, Bank for International Settlements, November.

Goodhart, Charles, 1988, The Evolution of Central Banks, Cambridge, MA: MIT Press.

Goodhart, Charles, Boris Hofmann, and Miguel Segoviano, 2004, Bank regulation and macroeconomic fluctuations, Oxford Review of Economic Policy 20, 591–615.

Goodhart, Charles, and Dirk Schoenmaker, 1995a, Should the functions of monetary policy and banking supervision be separated? Oxford Economic Papers 47, 539–560.

Goodhart, Charles, and Dirk Schoenmaker, 1995b, Institutional separation between supervisory and monetary agencies, in Charles Goodhart, ed., The Central Bank and the Financial System, London: Macmillan, pp. 332–412.

Gordy, Michael, 2003, A risk-factor model foundation for ratings-based bank capital rules, Journal of Financial Intermediation 12, 199–232.

Gordy, Michael, and Bradley Howells, 2006, Procyclicality in Basel II: Can we treat the disease without killing the patient? Journal of Financial Intermediation 15, 395–417.

Gorton, Gary, and George Pennacchi, 1995, Banks and loan sales: Marketing nonmaketable assets, Journal of Monetary Economics 35, 389–411.

Goyal, Vidhan, 2005, Market discipline of bank risk: Evidence from subordinated debt contracts, Journal of Financial Intermediation 14, 318–350.

Greenbaum, Stuart, and Anjan Thakor, 2007, Contemporary Financial Intermediation, 2nd Edition, Burlington, MA: Academic Press.

Gual, Jordi, 1999, Deregulation, integration, and market structure in European banking, Journal of the Japanese and International Economics 13, 372–396.

Gunther, Jeffery, and Robert Moore, 2003, Loss underreporting and the auditing role of bank exams, Journal of Financial Intermediation 12, 153–177.

Guthrie, Graeme, and Julian Wright, 2007, Competing payment schemes, Journal of Industrial Economics 55, 37–67.

Hadlock, Charles, Joel Houston, and Michael Ryngaert, 1999, The role of managerial incentives in bank acquisitions, Journal of Banking and Finance 23, 221–249.

Hagendorff, Jens, Michael Collins, and Kevin Keasey, 2008, Regulation, board monitoring, and merger performance: Evidence from acquiring banks in the U.S. and Europe, Manuscript, Leeds University Business School and University of Leeds, August 19.

Hall, Robert, 2007, Book review: Lectures on antitrust economics, Journal of Economic Literature 45, 1066–1070.

Hamalainen, Paul, Maximilian Hall, and Barry Howcroft, 2003, Market discipline: A theoretical framework for regulatory policy development, in George Kaufman, ed., Market Discipline in Banking: Theory and Evidence, Amsterdam: Elsevier, pp. 57–95.

Hancock, Diana, 1991, A Theory of Production for the Financial Firm, Boston: Kluwer.

Hancock, Diana, 1985a, Bank profitability, interest rates, and monetary policy, Journal of Money, Credit, and Banking 17, 189–202.

Hancock, Diana, 1985b, The financial firm: Production with monetary and nonmonetary goods, Journal of Political Economy 93, 859–880.

Hancock, Diana, and Myron Kwast, 2001, Using subordinated debt to monitor bank holding companies: Is it feasible? Journal of Financial Services Research 20, 147–187.

Hancock, Diana, and James Wilcox, 1994, Bank capital and the credit crunch: The roles of risk-weighted and unweighted capital regulation, Journal of the American Real Estate and Urban Economics Association 22, 59–94.

Hannan, Timothy, 2008, Consumer switching costs and firm pricing: Evidence from bank pricing of deposit accounts, Federal Reserve Economics and Discussion Series No. 2008-32, May 12.

Hannan, Timothy, 2007, ATM surcharge bans and bank market structure: The case of Iowa and its neighbors, Journal of Banking and Finance 31, 1061–1082.

Hannan, Timothy, 2006, Retail deposit fees and multimarket banking, Journal of Banking and Finance 30, 2561–2578.

Hannan, Timothy, 1997, Market share inequality, the number of competitors, and the HHI: An examination of bank pricing, Review of Industrial Organization 12, 23–35.

Hannan, Timothy, 1991, Foundations of the structure-conduct-performance paradigm in banking, Journal of Money, Credit, and Banking 23, 68–84.

Hannan, Timothy, and J. Nellie Liang, 1993, Inferring market power from time-series data: The case of the banking firm, International Journal of Industrial Organization 11, 205–218.

Hannan, Timothy, and Steven Pilloff, 2006, Acquisition targets and motives in the banking industry, Unpublished Manuscript, Board of Governors of the Federal Reserve System, September 19.

Hannan, Timothy, and Robin Prager, 2009, The profitability of small single-market banks in an era of multi-market banking, Journal of Banking and Finance 33, 263–271.

Hannan, Timothy, and Robin Prager, 2004, The competitive implications of multimarket bank branching, Journal of Banking and Finance 28, 1889–1914.

Hannan, Timothy, and Robin Prager, 1998, The relaxation of entry barriers in the banking industry: An empirical investigation, Journal of Financial Services Research 14, 171–188.

Hannan, Timothy, and Stephen Rhoades, 1987, Acquisition targets and motives: The case of the banking industry, Review of Economics and Statistics 69, 67–74.

Hardy, Daniel, 2006, Regulatory capture in banking, International Monetary Fund Working Paper WP/06/34, January.

Hasan, Iftekhar, William Curt Hunter, and Roswell Mathis III, 2002, Promotional expenditures, market competition, and thrift behavior, Journal of Business Research 50, 177–184.

Hasan, Iftekhar, and Stephen Smith, 1997, A note on fixed costs and the profitability of depository institutions, Managerial and Decision Economics 18, 47–53.

Haubrich, Joseph, and Paul Wachtel, 1993, Capital requirements and shifts in commercial bank portfolios, Federal Reserve Bank of Cleveland Economic Review 29 (Quarter 1), 1–15.

Hauswald, Robert, and Robert Marquez, 2006, Competition and strategic information acquisition in credit markets, Review of Financial Studies 19, 967–1000.

Hayashi, Fumiko, 2006, A puzzle of card payment pricing: Why are merchants still accepting card payments? Review of Network Economics 5, 144–174.

Heffernan, Shelagh, 1997, Modeling British interest rate adjustment: An error correction approach, Economica 64, 211–231.

Heggestad, Arnold, and John Mingo, 1976, Prices, nonprices, and concentration in commercial banking, Journal of Money, Credit, and Banking 8, 107–117.

Heitfield, Erik, and Robin Prager, 2004, The geographic scope of retail deposit markets, Journal of Financial Intermediation 25, 37–55.

Hellmann, Thomas, Kevin Murdock, and Joseph Stiglitz, 2000, Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough? American Economic Review 90, 147–165.

Hendricks, Darryll, and Beverly Hirtle, 1997, Bank capital requirements for market risk: The internal models approach, Federal Reserve Bank of New York Economic Policy Review, December, 3, 1–12.

Hernando, Ignacio, María Nieto, and Larry Wall, 2009, Determinants of domestic and cross-border bank acquisitions in the European Union, Journal of Banking and Finance 33, 1022–1032.

Herring, Richard, 2004, How can the invisible hand strengthen prudential regulation and how can prudential regulation strengthen the invisible hand? in Claudio Borio, William Hunter, George Kaufman, and Kostas Tsatsaronis, eds., Market Discipline Across Countries and Industries, Cambridge, MA: MIT Press, pp. 363–379.

Hester, Donald, and James Pierce, 1975, Bank Management and Portfolio Behavior, New Haven, CT: Yale University Press.

Hirtle, Beverly, 2007, The impact of network size on bank branch performance, Journal of Banking and Finance 31, 3782–3805.

Hirtle, Beverly, and Kevin Stiroh, 2007, The return to retail and the performance of U.S. banks, Journal of Banking and Finance 31, 1101–1133.

Hitt, Lorin, Frances Frei, and Patrick Harker, 1998, How financial firms decide on technology, Wharton Financial Institutions Center Working Paper No. 98-34-B, November.

Hodgman, Donald, 1963, Commercial Bank Loan and Investment Policy, Champaign, IL: Bureau of Economic and Business Research.

Hofmann, Boris, and Paul Mizen, 2004, Interest rate pass-through and monetary transmission: Evidence from individual financial institutions' retial rates, Economica 71, 99–123.

Holmström, Bengt, and Jean Tirole, 1997, Financial intermediation, loanable funds, and the real sector, Quarterly Journal of Economics 112, 663–691.

Holthausen, Cornelia, and Jean-Charles Rochet, 2006, Efficient pricing of large value interbank payment systems, Journal of Money, Credit, and Banking 38, 1797–1818.

Holthausen, Cornelia, and Thomas Rønde, 2005, Cooperation in international banking supervision, European Central Bank, CEPR Discussion Paper No. 4990, February 9.

Homölle, Susanne, 2004, Bank capital regulation, asset risk, and subordinated uninsured debt, Journal of Economics and Business 56, 443–468.

Hoshi, Takeo, Anil Kashyap, and David Scharfstein, 1991, Corporate structure, liquidity, and investment: Evidence from Japanese industrial groups, Quarterly Journal of Economics 106, 33–60.

Hotelling, Harold, 1929, Stability in competition, Economic Journal 39, 41–57.

Houston, Joel, Christopher James, and Michael Ryngaert, 2001, Where do merger games come from? Bank mergers from the perspective of insiders and outsiders, Journal of Financial Economics 60, 285–331.

Hovakimian, Armen, Edward Kane, and Luc Laeven, 2003, How country and safety-net characteristics affect bank risk-shifting, Journal of Financial Services Research 23, 177–204.

Huang, Rocco, and Lev Ratnovski, 2008, The dark side of bank wholesale funding, Federal Reserve Bank of Philadelphia Working Paper No. 09-3, November.

Hughes, Joseph, William Lang, Loretta Mester, and Choon-Geol Moon, 1999, The dollars and sense of bank consolidation, Journal of Banking and Finance 23, 291–324.

Hughes, Joseph, and Loretta Mester, 1998, Bank capitalization and cost: Evidence of scale economies in risk management and signaling, Review of Economics and Statistics 80, 314–325.

Hughes, Joseph, Loretta Mester, and Choon-Geol Moon, 2001, Are scale economies in banking elusive or illusive? Evidence obtained by incorporating capital structure and risk-taking into models of bank production, Journal of Banking and Finance 25, 2169–2208.

Hülsewig, Oliver Eric Mayer, and Timo Wollmershäuser, 2006, Bank loan supply and monetary policy transmission in Germany: An assessment based on matching impulse responses, Journal of Banking and Finance 30, 2893–2910.

Humala-Acuna, Alberto, 2005, Regime switching in interest rate pass-through and dynamic bank modeling with risks, Unpublished Manuscript, University of Warwick.

Humphrey, David, and Bent Vale, 2004, Scale economies, bank mergers, and electronic payments: A spline function approach, Journal of Banking and Finance 28, 1671–1696.

Hunter, William Curt, and Stephen Timme, 1995, Core deposits and physical capital: A reexamination of bank scale economies and efficiency with quasi-fixed inputs, Journal of Money, Credit, and Banking 27, 165–185.

Iannotta, Giuliano, 2006, Testing for opaqueness in the European banking industry: Evidence from bond credit ratings, Journal of Financial Services Research 30, 287–309.

Ioannidou, Vasso, 2005, Does monetary policy affect the central bank's role in bank supervision? Journal of Financial Intermediation 14, 58–85.

Iyer, Rajkamal, and Manju Puri, 2008, Understanding bank runs: The importance of depositor-bank relationships and networks, NBER Working Paper 14280, August.

Jacklin, Charles, 1987, Demand deposits, trading restrictions, and risk sharing, in Edward Prescott and Neil Wallace, eds., Contractual Arrangements for Intertemporal Trade, Minneapolis: University of Minnesota Press.

Jackson, Patricia, Craig Furfine, Hans Groeneveld, Diana Hancock, David Jones, William Perraudin, Lawrence Radecki, and Masao Yoneyama, 1999, Capital requirements and bank behavior: The impact of the Basel accord, Basel Committee on Banking Supervision, Working Paper No. 1, April.

Jackson, William III, 1997, Market structure and the speed of price adjustments: Evidence of non-monotonicity, Review of Industrial Organization 12, 37–57.

Jackson, William, 1992, Is the market well defined in bank merger and acquisition analysis? Review of Economics and Statistics 74, 655–661.

Jacques, Kevin, and Peter Nigro, 1997, Risk-based capital, portfolio risk, and bank capital: A simultaneous equations approach, Journal of Economics and Business 49, 533–547.

James, Christopher, 1987, Some evidence on the uniqueness of bank loans, Journal of Financial Economics 19, 217–235.

Janicki, Hubert, and Edward Simpson Prescott, 2006, Changes in the size distribution of U.S. banks: 1960–2005, Federal Reserve Bank of Richmond Economic Quarterly 92, 291–316.

Jayaratne, Jith, and Philip Strahan, 1998, Entry restrictions, industry evolution, and dynamic efficiency: Evidence from commercial banking, Journal of Law and Economics 41, 239–274.

Jeitschko, Thomas, and Shin Dong Jeung, 2005, Incentives for risk-taking in banking: A unified approach. Journal of Banking and Finance 29, 759–777.

Jones, Kenneth, and Tim Critchfield, 2008, Consolidation in the U.S. banking industry: Is the "long, strange trip" about to end?, in Anjan Thakor and Arnoud Boot, eds., Handbook of Financial Intermediation and Banking, Amsterdam: Elsevier, pp. 309–346.

Jordan, John, Joe Peek, and Eric Rosengren, 2000, The market reaction to the disclosure of supervisory actions: Implications for bank transparency, Journal of Financial Intermediation 9, 298–319.

Kahane, Yehuda, 1977, Capital adequacy and the regulation of financial intermediaries, Journal of Banking and Finance 1, 207–218.

Kahn, Charles, George Pennacchi, and Ben Spranzetti, 2005, Bank consolidation and consumer loan rates, Journal of Business 78, 99–133.

Kahn, Charles, George Pennacchi, and Ben Spranzetti, 1999, Bank deposit rate clustering: Theory and empirical evidence, Journal of Finance 6, 2185–2214.

Kahn, Charles, and William Roberds, 2009, Why pay? An introduction to payments economics, Journal of Financial Intermediation 18, 1–23.
Kane, Edward, 2000, Incentives for banking megamergers: What motives might regulators infer from event-study evidence? Journal of Money, Credit, and Banking 32, 671–701.
Kane, Edward, 1989, The S&L Insurance Mess: How Did It Happen?, Washington, DC: Urban Institute Press.
Kane, Edward, 1985, The Gathering Crisis in Federal Deposit Insurance, Cambridge, MA: MIT Press.
Karaoglu, N. Emre, 2005, Regulatory capital and earnings management in banks: The case of loan sales and securitizations, FDIC Center for Financial Research Working Paper No. 2005-5, May.
Karceski, Jason, Steven Ongena, and David Smith, 2005, The impact of bank consolidation on commercial borrower welfare, Journal of Finance 60, 2043–2082.
Kashyap, Anil, and Jeremy Stein, 2004, Cyclical implications of the Basel II capital standards, Federal Reserve Bank of Chicago Economic Perspectives, 18–31.
Kaufman, George, 2006, Basel II has been a costly distraction on the road to minimizing the societal cost of bank failures, Presented at the Sixth Annual FDIC Research Conference, September 13.
Kauko, Karlo, 2005, Bank interest rates in a small European economy: Some exploratory macro level analyses using Finnish data, Bank of Finland Research Discussion Paper 9, 2005.
Keeley, Michael, 1990, Deposit insurance, risk, and market power in banking, American Economic Review 80, 1183–1200.
Keeley, Michael, 1988, Bank capital regulation in the 1980s: Effective or ineffective? Federal Reserve Bank of San Francisco Economic Review, Winter, 3–20.
Keeley, Michael, and Frederick Furlong, 1990, A reexamination of the mean-variance analysis of bank capital regulation, Journal of Banking and Finance 14, 69–84.
Kim, Daesik, and Anthony Santomero, 1988. Risk in banking and capital regulation, Journal of Finance 43, 1219–1233.
Kim, Moshe, Doron Kliger, and Bent Vale, 2003, Estimating switching costs: The case of banking, Journal of Financial Intermediation 12, 25–56.
Kim, Moshe, Eirik Gaard Kristiansen, and Bent Vale, 2005, Endogenous product differentiation in credit markets: What do borrowers pay for? Journal of Banking and Finance 29, 681–699.
Kim, Moshe, and Bent Vale, 2001, Non-price strategic behavior: The cost of bank branches, International Journal of Industrial Organization 19, 1583–1602.
King, Thomas, 2008, Discipline and liquidity in the interbank market, Journal of Money, Credit, and Banking 40, 295–317.
Kiser, Elizabeth, 2002a, Household switching behavior at depository institutions: Evidence from survey data, Antitrust Bulletin 47, 619–640.
Kiser, Elizabeth, 2002b, Predicting household switching behavior and switching costs at depository institutions, Review of Industrial Organization 20, 349–365.
Klein, Benjamin, 1974, Competitive interest payments on bank deposits and the long-run demand for money, American Economic Review 64, 931–949.
Klein, Michael, 1978, The implicit deposit rate: Issues and applications, Federal Reserve Bank of Richmond Economic Review 64, 3–12.
Klein, Michael, 1971, A theory of the banking firm, Journal of Money, Credit, and Banking 3, 205–218.
Kahn, Charles, George Pannacchi, and Ben Sopranzetti, 2005, Bank consolidation and the dynamics of consumer loan interest rates, Journal of Business 78, 99–133.
Kocherlakota, Narayana, and Ilhyock Shim, 2007, Forbearance and prompt corrective action, Journal of Money, Credit, and Banking 39, 1107–1129.
Koehn, Michael and Anthony Santomero, 1980, Regulation of bank capital and portfolio risk. Journal of Finance 35, 1235–1244.
Koetter, Michael, J.W.B. Bos, F. Heid, J.W. Kolari, C.J.M. Kool, and D. Porath, 2007, Accounting for distress in bank mergers, Journal of Banking and Finance 31, 3200–3217.

Kohers, Theodor, and W. Gary Simpson, 1981, Concentration and advertising in the United States savings and loan industry, Applied Economics 13, 79–88.

Kopecky, Kenneth, and David VanHoose, 2009, Imperfect competition in bank retail markets, deposit and loan rate dynamics, and incomplete pass through, Unpublished Manuscript, Temple University and Baylor University.

Kopecky, Kenneth, and David VanHoose, 2006, Capital regulation, heterogeneous monitoring costs, and aggregate loan quality, Journal of Banking and Finance 30 (August), 2235–2255.

Kopecky, Kenneth, and David VanHoose, 2004a, Bank capital requirements and the monetary transmission mechanism, Journal of Macroeconomics 26, 443–464.

Kopecky, Kenneth, and David VanHoose, 2004b, A model of the monetary sector with and without binding capital requirements, Journal of Banking and Finance 28, 633–646.

Koskela, Erkki, and Rune Stenbacka, 2000, Is there a tradeoff between bank competition and financial fragility? Journal of Banking and Finance 24, 1853–1873.

Kroszner, Randall, and Philip Strahan, 2001, Obstacles to optimal policy: The interplay of politics and economics in shaping bank supervision and regulatory reforms, in Frederic Mishkin, ed., Prudential Supervision: What Works and What Doesn't, Chicago: University of Chicago Press, pp. 233–272.

Kroszner, Randall, and Philip Strahan, 1999, What drives deregulation? Economics and politics of the relaxation of bank branching restrictions, Quarterly Journal of Economics 114, 1437–1467.

Kwan, Simon, 2001, Financial modernization and banking theories, Federal Reserve Bank of San Francisco Economic Letter, No. 2001-37, December 21.

Kwast, Myron, Daniel Covitz, Diana Hancock, James Houpt, David Adkins, Norah Barger, Barbary Bouchard, John Connolly, Thomas Brady, William English, Douglas Evanoff, and Larry Wall, 1999, Using subordinated debt as an instrument of market discipline, Staff Study 172, Board of Governors of the Federal Reserve System.

Kwast, Myron, and S. Wayne Passmore, 2000. The subsidy provided by the federal safety net: Theory and evidence, Journal of Financial Services Research 17, 125–145.

Kwast, Myron, Martha Starr-McCluer, and John Wolken, 1997, Market definition and the analysis of antitrust in banking, Antitrust Bulletin 42, 973–995.

Lacker, Jeffrey, and John Weinberg, 2003, Payment economics: Studying the economics of exchange, Journal of Monetary Economics 50 (2), 381–387.

Laderman, Elizabeth, 2008, Small business lending and banking competition, Federal Reserve Bank of San Francisco Economic Letter No. 2008-15, May 9.

Laderman, Elizabeth, 2007, market power and relationships in small business lending, Federal Reserve Bank of San Francisco Working Paper 2007-07.

Laffont, Jean-Jacques, and Jean Tirole, 1991, The politics of government decision-making: A theory of regulatory capture, Quarterly Journal of Economics 106, 1089–1127.

Lam, Chun, and Andrew Chen, 1985, Joint effects of interest rate deregulation and capital requirements on optimal bank portfolio adjustments, Journal of Finance 40, 563–575.

Landskroner, Yoram, and Jacob Paroush, 2008, Bank management and market discipline, Journal of Economics and Business 60, 395–414.

Landsman, Wayne, 2006, Fair value accounting for financial instruments: Some implications for bank regulation, BIS Working Paper No. 209, August.

Lane, Timothy, 1993, Market discipline, IMF Staff Papers 40, 53–88.

Lang, William, Loretta Mester, and Todd Vermilyea, 2008, Competitive effects of Basel II on U.S. bank credit card lending, Journal of Financial Intermediation 17, 478–508.

Lang, William, and Douglas Robertson, 2002, Analysis of proposals for a minimum subordinated debt requirement, Journal of bEconomics and Business 54, 115–136.

Lapp, John, 1976, Market structure and advertising in the savings and loan industry, Review of Economics and Statistics 58, 202–208.

Lee, Kwang-Won, and Ian Sharpe, 2009, Does a bank's loan screening and monitoring matter? Journal of Financial Services Research 35, 33–52.

Liang, Nellie, and Stephen Rhoades, 1988, Geographic diversification and risk in banking, Journal of Economics and Business 40, 271–284.

Llewellyn, David, and David Mayes, 2003, The role of market discipline in handling problem banks, in George Kaufman, ed., Market Discipline in Banking: Theory and Evidence, Amsterdam: Elsevier, pp. 183–210.

Lown, Cara, and Stavros Peristiani, 1996, The behavior of consumer loan rates during the 1990 credit slowdown, Journal of Banking and Finance 20, 1673–1694.

Lown, Cara, and John Wood, 2003, The determination of commercial bank reserve requirements, Review of Financial Economics 12, 83–98.

Maclachlan, Fiona, 2001, Market discipline in bank regulation: Panacea or paradox? The Independent Review 6, 227–234.

Magalhaes, Romulo, and Josep Tribó, 2009, "Bank Regulations and Loan Contracts," Working Paper, Central Bank of Brazil and Univeridad Carlos III de Madrid, February.

Mailath, George, and Loretta Mester, 1994, A positive analysis of bank closure, Journal of Financial Intermediation 3, 272–299.

Mallett, Ted, and Anindya Sen, 2001, Does local competition impact interest rates charges on small business loans? Empirical evidence from Canada, Review of Industrial Organization 19, 437–452.

Marinč, Matej, 2008, Essays on Bank Monitoring, Regulation, and Competition, Ph.D. Dissertation, University of Amsterdam.

Marini, François, 2003, Bank insolvency, deposit insurance, and capital adequacy, Journal of Financial Services Research 24, 67–78.

Marotta, Giuseppe, 2007, Structural breaks in the interest rate pass-through and the euro: A cross-country study in the euro area and the UK, Manuscript, University of Modena and CEFIN.

Marshall, David, and Edward Simpson Prescott, 2006, State-contingent bank regulation with unobserved actions and unobserved characteristics, Journal of Economic Dynamics and Control 30, 2015–2049.

Marshall, David, and Edward Simpson Prescott, 2001, Bank capital regulation with and without state-contingent penalties, Carnegie-Rochester Series on Public Policy 54, 139–184.

Martiez-Miera, David, and Rafael Repullo, 2008, Does competition reduce the risk of bank failure? Centre for Economic Policy Research Discussion Paper No. 6669, January.

Martín-Oliver, Alfredo, 2009 Market power in multiple-choice demand functions of banking services: An application to Spanish banks, Presented at International Industrial Organization Society Conference, Boston, MA, April.

Martín-Oliver, Alfredo, and Vicente Sals-Fumás, 2008, The output and profit contribution of information technology and advertising investment in banks, Journal of Financial Intermediation 17, 229–255.

Martín-Oliver, Alfredo, Vicente Sals-Fumás, and Jesús Saurina, 2008, Search cost and price dispersion in vertically related markets: The case of bank loans and deposits, Review of Industrial Organization 33, 297–323.

Masciandaro, Donato, and Marc Quintyn, 2008, Helping hand or grabbing hand? Politicians, supervision regime, financial structure and market view, North American Journal of Economics and Finance 19, 153–173.

Matthews, Kent, and John Thompson, 2005, The Economics of Banking, Chichester, West Sussex: Wiley.

Matutes, Carmen, and Xavier Vives, 2000, Imperfect competition, risk taking, and regulation in banking, European Economic Review 44, 1–34.

Matutes, Carmen, and Xavier Vives, 1996, Competition for deposits, fragility, and insurance, Journal of Financial Intermediation 5, 184–216.

McDill, Kathleen, and Andrea Maechler, 2003, Do uninsured depositors vote with their feet? in George Kaufman, ed., Market Discipline in Banking: Theory and Evidence, Amsterdam: Elsevier, pp. 211–248.

McIntosh, James, 2002, A welfare analysis of Canadian chartered bank mergers, Canadian Journal of Economics 35, 457–475.

Mercieca, Steve, Klaus Schaeck, and Simon Wolfe, 2007, Small European banks: Benefits from diversification? Journal of Banking and Finance, 1975–1998.
Merris, Randall, 1985, Explicit interest and implicit demand deposit interest, Journal of Money, Credit, and Banking 17, 528–533.
Mester, Loretta, 2008, Optimal industrial structure in banking, in Anjan Thakor and Arnoud Boot, eds., Handbook of Financial Intermediation and Banking, Amsterdam: Elsevier, pp. 133–162.
Mester, Loretta, Leonard Nakamura, and Micheline Renault, 2005, Transactions accounts and loan monitoring, Research Department, Federal Reserve Bank of Philadelphia, Working Paper No. 05-14, June.
Miller, Roger Leroy, 2009, Economics Today, Boston: Addison-Wesley.
Miller, Stephen, 1975, A theory of the banking firm: Comment, Journal of Monetary Economics 1, 123–128.
Milne, Alistair, 2002, Bank capital regulation as an incentive mechanism: Implications for portfolio choice, Journal of Banking and Finance 26, 1–23.
Mingo, John. 1975. Regulatory influence on bank capital investment. Journal of Finance 30, 1111–1121.
Mishkin, Frederic, 2006, How big a problem is too big to fail? A review of Gary Stern and Ron Feldman's too big to fail: The hazards of bank bailouts. Journal of Economic Literature 44, 988–1004.
Mishkin, Frederic, 1999, Financial consolidation: Dangers and opportunities, Journal of Banking and Finance 23, 675–691.
Mitchell, Douglas, 1979, Explicit and implicit demand deposit interest: Substitutes or complements from the bank's point of view, Journal of Money, Credit, and Banking 11, 182–191.
Mlima, Aziz Ponary, and Lennart Hjalmarsson, 2002, Measurement of inputs and outputs in the banking industry, Tanzanet Journal 3, 12–22.
Molnár, József, 2008, Market power and merger simulation in retail banking, Bank of Finland Research Discussion Paper #4-2008.
Morgan, Donald, 2002, Rating banks: Risk and uncertainty in an opaque industry, American Economic Review 92, 874–888.
Morgan, Donald, and Kevin Stiroh, 2001, Market discipline of banks: The asset test, Journal of Financial Services Research 20, 195–208.
Morrison, Alan, and Lucy White, 2005. Crises and capital requirements in banking, American Economic Review 95, 1548–1572.
Nelson, Richard, 1985, Branching, scale economies, and banking costs, Journal of Banking and Finance 9, 177–191.
Neuberger, Doris, 1998, Industrial organization of banking: A review, International Journal of the Economics of Business 5, 97–118.
Neumark, David and Steven Sharpe, 1992, Market structure and the nature of price rigidity: Evidence from the market for consumer deposits, Quarterly Journal of Economics 107, 657–680.
Neven, Damien, and Lars-Hendrik Röller, 1999, An aggregate structural model of competition in the European banking industry, International Journal of Industrial Organization 17, 1059–1074.
Nier, Erlend, and Ursel Baumann, 2006, Market discipline, disclosure, and moral hazard in banking, Journal of Financial Intermediation 15, 332–361.
Niu, Jijun, 2008, Bank competition, risk, and subordinated debt, Journal of Financial Services Research 33, 37–56.
Office of the Comptroller of the Currency, 2006a, Regulatory impact analysis for risk-based capital guidelines, capital adequacy guidelines, capital maintenance, and domestic capital modifications, International and Economic Affairs.
Office of the Comptroller of the Currency, 2006b, Regulatory impact analysis for risk-based capital standards, revised capital adequacy guidelines, Basel II: Advanced Approach, International and Economic Affairs.

Ogura, Yoshiaki, 2006, Learning from a rival bank and lending boom, Journal of Financial Intermediation 15, 535–555.

O'Hara, Maureen, and Wayne Shaw, 1990, Deposit insurance and wealth effects: The value of banking 'too big to fail,' Journal of Finance 45, 1587–1600.

Ongena, Steven, and Alexander Popov, 2009, Interbank market integration, bank competition, and loan rates, Working Paper, Tilburg University and the European Central Bank, February.

Ongena, Steven, and David Smith, 2001, The duration of bank relationships, Journal of Financial Economics 61, 449–475.

Örs, Evren, 2006, The role of advertising in commercial banking, Centre for Economic Policy Research Discussion Paper No. 5461.

Östberg, Per, 2006, Disclosure, investment, and regulation, Journal of Financial Intermediation 15, 285–306.

Palia, Darius, 1993, The managerial, regulatory, and financial determinants of bank merger premiums, Journal of Industrial Economics 41, 91–102.

Panzar, John C., and James N. Rosse, 1987, Testing for 'monopoly' equilibrium, Journal of Industrial Economics 35, 443–456.

Park, Kwangwoo, and George Pennacchi, 2008, Harming depositors and helping borrowers: The disparate impact of bank consolidation, Review of Financial Studies, 22, 1–40.

Park, Sangkyun, and Stavros Peristiani, 2007, Are bank shareholders enemies of regulators or a potential source of market discipline? Journal of Banking and Finance 31, 2493–2515.

Park, Sangkyun, and Stavros Peristiani, 1998, Market discipline by thrift depositors, Journal of Money, Credit, and Banking 30, 347–364.

Pasiouras, Fotios, Sailesh Tanna, and Constantin Zopounidis, 2008, Banking regulations, cost and profit efficiency: Cross-country evidence, Manuscript, University of Bath, Coventry University, and Technical University of Crete.

Pasiouras, Fotios, Sailesh Tanna, and Chrysovalantis Gaganis, 2007, What drives acquisitions in the EU banking industry? Coventry University, Economics, Finance, and Accounting Applied Research Working Paper No. 2007-3.

Pattison, John, 2006, International financial cooperation and the number of adherents: The Basel committee and capital regulation, Open Economies Review 17, 443–458.

Pederzoli, Chirara, and Constanza Torricelli, 2005, Capital requirements and business cycle regimes: Forward-looking modeling of default probabilities, Journal of Banking and Finance 29, 3121–3140.

Peek, Joe and Eric Rosengren, 1995a, The capital crunch: Neither a borrower nor a lender be, Journal of Money, Credit, and Banking 27, 625–638.

Peek, Joe and Eric Rosengren, 1995b, Bank regulation and the credit crunch, Journal of Banking and Finance 19, 679–692.

Peek, Joe, Eric Rosengren, and Geoffrey Tootell, 2003, Does the federal reserve possess an exploitable informational advantage? Journal of Monetary Economics 50, 817–839.

Peek, Joe, Eric Rosengren, and Geoffrey Tootell, 2001, Synergies between bank supervision and monetary policy, in Frederic Mishkin, ed., Prudential Supervision: What Works and What Doesn't, Chicago: University of Chicago Press, pp. 273–293.

Peek, Joe, Eric Rosengren, and Geoffrey Tootell, 1999, Is bank supervision central to central banking? Quarterly Journal of Economics 114, 629–653.

Peltzman, Sam, 1976, Toward a more general theory of regulation, Journal of Law and Economics 19, 211–240.

Peltzman, Sam, 1970, Capital regulation in commercial banking and its relation to portfolio regulation, Journal of Political Economy 78, 1–26.

Pennacchi, George, 2005, Risk-based capital standards, deposit insurance, and procyclicality, Journal of Financial Intermediation 14, 432–465.

Peresetsky, Anatoly, Alexander Karminsky, and Sergei Golovan, 2007, Russian banks' private deposit interest rates and market discipline, Bank of Finland BOFIT Discussion Paper No. 2/2007, January 23.

Perloff, Jeffrey, Larry Karp, and Amos Golan, 2007, Estimating market power and strategies, Cambridge: Cambridge University Press.

Peters, Craig, 2003, Evaluating the performance of merger simulation: Evidence from the U.S. airline industry, Northwestern University Center for the Study of Industrial Organization Working Paper #0032, January.

Petersen, Mitchell, and Raghuram Rajan, 1995, The effect of credit market competition on lending relationships, Quarterly Journal of Economics 110, 407–443.

Petersen, Mitchell, and Raghuram Rajan, 1994, The benefits of lending relationships: Evidence from small business data, Journal of Finance 49, 3–37.

Pilloff, Steven, 2009, The banking industry, in James Brock, ed., The Structure of American Industry, 12th Edition, Upper Saddle River, NJ: Pearson.

Pilloff, Steven, 1999, Does the presence of big banks influence competition in local markets? Journal of Financial Intermediation Research 15, 159–177.

Podpiera, Richard, 2006, Does compliance with Basel core principles bring any measurable benefits? IMF Staff Papers 53, 306–326.

Posner, Richard, 1974, Theories of economic regulation, Bell Journal of Economics and Management Science 5, 335–358.

Posner, Richard, 1971, Taxation by regulation, Bell Journal of Economics and Management Science 2, 22–50.

Presbitero, Andrea, and Alberto Zazzaro, 2009, Competition and relationship lending: Friends or foes? Money and Finance Research Group Working Paper No. 13, Università Politechnica delle Marche, January.

Prescott, Edward Simpson, 2008, Should bank supervisors disclose information about their banks? Federal Reserve Bank of Richmond Economic Quarterly 94, 1–16.

Pringle, John, 1973, A theory of the banking firm: A comment, Journal of Money, Credit, and Banking 5, 990–996.

Radecki, Lawrence, 1999, Banks' payment-driven revenues, Federal Reserve Bank of New York Economic Policy Review 5, 53–70.

Radecki, Lawrence, 1998, The expanding geographic reach of retail banking markets, Federal Reserve Bank of New York Economic Policy Review 4, 15–34.

Ratti, Ronald, Sunglyong Lee, and Youn Seol, 2008, Bank concentration and financial constraints on firm-level investment in Europe, Journal of Banking and Finance 32, 2684–2694.

Repullo, Rafael, 2004, Capital requirements, market power, and risk-taking in banking, Journal of Financial Intermediation 13, 156–182.

Repullo, Rafael, and Javier Suarez, 2004, Loan pricing under Basel capital requirements, Journal of Financial Intermediation 13, 496–521.

Rhoades, Stephen, 2000a, Bank mergers and banking structure in the United States, 1980–98, Staff Study 174, Board of Governors of the Federal Reserve System, August.

Rhoades, Stephen, 2000b, Retail commercial banking: An update on a period of extraordinary change, Review of Industrial Organization 16, 357–366.

Rhoades, Stephen, 1997, Have barriers to entry in retail commercial banking disappeared? Antitrust Bulletin 42, 997–1013.

Rhoades, Stephen, 1994, 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, July.

Rhoades, Stephen, 1982, Welfare loss, redistribution effects, and restriction of output due to monopoly in banking, Journal of Monetary Economics 9, 375–387.

Rhoades, Stephen, 1980, Monopoly and expense preference behavior: An empirical investigation of the behavioralist hypothesis, Southern Economic Journal 47, 419–432.

Richards, Timothy, Ram Acharya, and Albert Kagan, 2008, Spatial competition and market power in banking, Journal of Economics and Business 60, 436–454.

Rime, Bertrand, 2001, Capital requirements and bank behavior: Empirical evidence for Switzerland, Journal of Banking and Finance 25, 789–805.

Risk Transfer Magazine, 2005, Basel costs don't dazzle bankers. September 26. (http://www.risktransfermagazine.com/xq/asp/sid.0/articleid.AF7CFB58-91C1-4FE4-9649-2F2298A9A292/qx/display.htm)

Rivard, Richard, and Christopher Thomas, 1997, The effect of interstate banking on large bank holding company profitability and risk, Journal of Economics and Business 49, 61–76.

Rochet, Jean-Charles, 2008, Why are there so many banking crises? in Jean-Charles Rochet, ed., Why Are There So Many Banking Crises: The Politics and Policy of Bank Regulation, Princeton, NJ: Princeton University Press, pp. 22–34.

Rochet, Jean-Charles, 1992, Capital requirements and the behavior of commercial banks, European Economic Review 36, 1137–1178.

Rochet, Jean-Charles, and Jean Tirole, 2008, Competition policy in two-sided markets, with a special emphasis on payment cards, in Paulo Buccirossi, ed., Handbook of Antitrust Economics, Cambridge, MA: MIT Press, pp. 543–582.

Rochet, Jean-Charles, and Jean Tirole, 2006a, Two-sided markets: A progress report, Rand Journal of Economics 37, 645–667.

Rochet, Jean-Charles, and Jean Tirole, 2006b, Externalities and regulation in card payment systems, Review of Network Economics 5, 1–25.

Rochet, Jean-Charles, and Jean Tirole, 2006c, Tying in two-sided markets and the honor all cards rule, Working Paper, Toulouse University, February 9.

Rochet, Jean-Charles, and Jean Tirole, 2003, Platform competition in two-sided markets, Journal of the European Economic Association 1, 990–1029.

Rochet, Jean-Charles, and Jean Tirole, 2002, Cooperation among competitors: Some economics of payment card associations, Rand Journal of Economics 33, 549–570.

Rochet, Jean-Charles, and Xavier Vives, 2004, Coordination failures and the lender of last resort: Was Bagehot right after all? Journal of the European Economic Association 2, 1116–1147.

Rosen, Richard, 2007, Bank market conditions and deposit interest rates, Journal of Banking and Finance 31, 3862–3884.

Rosen, Richard, 2003, Is three a crowd? Competition among regulators in banking, Journal of Money, Credit, and Banking 35, 967–998.

Rosenbluth, Frances, and Ross Schaap, 2003, The domestic politics of banking regulation, International Organization 57, 307–336.

Salop, Steven, 1979, Monopolistic competition with outside goods, Bell Journal of Economics 10, 141–156.

Salop, Steven, and David Scheffman, 1987, Cost-raising strategies, Journal of Industrial Economics 36, 19–34.

Salop, Steven, and David Scheffman, 1983, Raising rivals' costs, American Economic Review Papers and Proceedings 73, 267–271.

Sander, Harald, and Stefanie Kleimeier, 2006, Convergence of interest rate pass-through in a wider euro zone? Economic Systems 30, 405–423.

Sander, Harald, and Stefanie Kleimeier, 2004, Convergence in euro-zone retail banking? What interest rate pass-through tells us about monetary policy transmission, competition, and integration, Journal of International Money and Finance 23, 461–492.

Santomero, Anthony, 1984, Modeling the banking firm, Journal of Money, Credit, and Banking 16, 576–602.

Santos, João, 2001, Bank capital regulation in contemporary banking theory: A review of the literature, Financial Markets, Institutions, and Instruments 10, 41–84.

Santos, João, 1999, Bank capital and equity investment regulations, Journal of Banking and Finance 23, 1095–1120.

Sapienza, Paola, 2002, The effects of banking mergers on loan contracts, Journal of Finance 57, 329–367.

Saunders, Anthony, 2001, Comments on Evanoff and Wall/Hancock and Kwast, Journal of Financial Services Research 20, 189–194.

Schaeck, Klaus, and Martin Čihák, 2008, How does competition affect efficiency and soundness in banking? European Central Bank Working Paper No. 932, September.
Schaeck, Klaus, Martin Čihák, and Simon Wolfe, 2006, Are more competitive banking systems more stable? IMF Working Paper WP/06/143.
Scheffman, David, and Richard Higgins, 2003, Twenty years of raising rivals' costs: History, assessment, and future, George Mason Law Review 12, 371–387.
Schmalensee, Richard, 2002, Payment systems and interchange fees, Journal of Industrial Economics 50, 103–122.
Schmid, Frank, 1994, Should bank branching be regulated? Theory and empirical evidence from four European countries, Journal of Regulatory Economics 6, 137–149.
Scholnick, Barry, 1999, Interest rate asymmetries in long-term loan and deposit markets, Journal of Financial Services Research 16, 5–26.
Scholnick, Barry, 1996, Asymmetric adjustment of commercial bank interest rates: Evidence from Malaysia and Singapore, Journal of International Money and Finance 15, 485–496.
Schwartz, Marius, and Daniel Vincent, 2006, The no surcharge rule and card user rebates: Vertical control by a payment network, Review of Network Economics 5, 72–102.
Scott, Jonathan, and William Dunkelberg, 2003, Bank mergers and small firm financing, Journal of Money, Credit, and Banking 35, 999–1017.
Scott, John, 1978, Non-price competition in banking markets, Southern Economic Journal 44, 594–605.
Sealey, C.W., Jr., 1985, Portfolio separation for stockholder-owned depository financial intermediaries, Journal of Banking and Finance 9, 477–490.
Sealey, C.W., Jr., 1977, A further reconsideration of optimal reserve management for depository financial institutions, Southern Economic Journal 44, 117–124.
Sealey, C.W., Jr., and James Lindley, 1977, Inputs, outputs, and a theory of production and cost at depository financial institutions, Journal of Finance 32, 1251–1266.
Shadow Financial Regulatory Committee, 2001, Statement on the Basel Committee's revised capital accord proposal, Statement No. 169, February 26.
Shaffer, Sherrill, 2004a, Patterns of competition in banking, Journal of Economics and Business 56, 287–313.
Shaffer, Sherrill, 2004b, Comment on "What drives bank competition? Some international evidence" by Stijn Claessens and Luc Laeven, Journal of Money, Credit, and Banking 36, 585–592.
Shaffer, Sherrill, 2002, Conduct in a banking monopoly, Review of Industrial Organization 20, 221–238.
Shaffer, Sherrill, 2001, Banking conduct before the European single banking license: A cross-country comparison, North American Journal of Economics and Finance 12, 79–104.
Shaffer, Sherrill, 1998, Functional forms and declining average costs, Journal of Financial Services Research 14, 91–115.
Shaffer, Sherrill, 1993, A test of competition in Canadian banking, Journal of Money, Credit, and Banking 25, 49–61.
Shaffer, Sherrill, 1989, Competition in the U.S. banking industry, Economics Letters 29, 321–323.
Shaffer, Sherrill, and Soumya Srinivasan, 2002, Structure-pricing linkages among single-market banks, controlling for credit quality, Applied Economics Letters 9, 653–656.
Sharpe, Steven, 1997, The effect of consumer switching costs on prices: A theory and its application to the bank deposit market, Review of Industrial Organization 12, 79–94.
Sharpe, Steven, 1990, Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships, Journal of Finance 49, 1069–1087.
Sheehan, Richard, 2006, Price stickiness in bank deposit rates: An analysis of intra-market relationships, Unpublished Manuscript, University of Notre Dame.
Shim, Ilhyock, 2006, Dynamic prudential regulation: Is prompt corrective action optimal? BIS Working Paper No. 206, May.

Shiman, Daniel, 2007, The intuition behind Sutton's theory of endogenous sunk costs, Working Paper, Federal Communications Commission.

Shrieves, Ronald, and Drew Dahl, 2003, Discretionary accounting and the behavior of Japanese banks under financial duress, Journal of Banking and Finance 27, 1219–1243.

Shrieves, Ronald, and Drew Dahl, 1992, The relationship between risk and capital at commercial banks, Journal of Banking and Finance 16, 439–457.

Shull, Bernard, and Gerald Hanweck, 2001, Bank Mergers in a Deregulated Environment: Promise and Peril, Westport, CT: Quorum Books.

Simons, Katerina, and Joanna Stavins, 1998, Has antitrust policy in banking become obsolete? Federal Reserve Bank of Boston New England Economic Review, March/April, 13–26.

Sironi, Andrea, 2001, An analysis of European banks' SND issues and its implications for the design of a mandatory subordinated debt policy, Journal of Financial Services Research 20, 223–266.

Sleet, Christopher, and Bruce Smith, 2000, Deposit insurance and lender-of-last-resort functions, Journal of Money, Credit, and Banking 32, 518–575.

Sloan, Steven. 2007, Basel II compliance costing millions, American Banker. April 20.

Smirlock, Michael, 1985, Evidence on the (non) relationship between concentration and profitability in banking, Journal of Money, Credit, and Banking 17, 69–83.

Smirlock, Michael, and David Brown, 1986, Collusion, efficiency, and pricing behavior: Evidence from the banking industry, Economic Inquiry 24, 85–96.

Smith, Bruce, 1984, Private information, deposit interest rates, and the 'stability' of the banking system, Journal of Monetary Economics 14, 293–317.

Sørensen, Christoffer, and Thomas Werner, 2006, Bank interest rate pass-through in the euro area: A cross-country comparison, European Central Bank Working Paper No. 580, January 2006.

Startz, Richard, 1983, Competition and interest rate ceilings in commercial banking, Quarterly Journal of Economics 98, 255–265.

Stern, Gary, and Ron Feldman, 2004, Too Big to Fail: The Hazards of Bank Bailouts, Washington, DC: Brookings Institution.

Stigler, George, 1971, The theory of economic regulation, Bell Journal of Economics and Management Science 2, 3–21.

Stigler, George, 1955, Introduction, in National Bureau of Economic Research, Business Concentration and Price Policy, Princeton, NJ: Princeton University Press.

Stigler, George, and Robert Sherwin, 1985, The extent of the market, Journal of Law and Economics 29, 555–585.

Stiroh, Kevin, 2004, Diversification in banking: Is noninterest income the answer? Journal of Money, Credit, and Banking 36, 853–882.

Stiroh, Kevin, and Philip Strahan, 2003, Competitive dynamics of deregulation: Evidence from U.S. banking, Journal of Money, Credit, and Banking 35, 801–828.

Stolz, Stéphanie, 2007, Bank Capital and Risk-Taking: The Impact of Capital Regulation, Charter Value, and the Business Cycle, Berlin: Springer.

Strahan, Philip, 2008, Bank structure and lending: What we do know and do not know, in Anjan Thakor and Arnoud Boot, eds., Handbook of Financial Intermediation and Banking, Amsterdam: Elsevier, pp. 97–132.

Sun, Mingchun, and Edison Tse, 2007, When does the winner take all in two-sided markets? Review of Network Economics 6, 16–40.

Sutton, John, 2001, Technology and Market Structure: Theory and History, Cambridge, MA: MIT Press.

Sutton, John, 1991, Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration. Cambridge, MA: MIT Press.

Swank, Job, 1996, Theories of the banking firm: A review of the literature, Bulletin of Economic Research 48, 173–207.

Szegö, Giorgio, 1980, Portfolio Theory with Application to Bank Asset Management, New York: Academic Press.

Tadesse, Solomon, 2005, Banking fragility and disclosure: International evidence, Manuscript, University of South Carolina.
Tarullo, Daniel, 2008, Banking on Basel: The Future of International Financial Regulation, Washington, DC: Peterson Institute for International Economics.
Taylor, John, 2009, Getting Off Track: How Government Actions and Interventions Cause, Prolonged, and Worsened the Financial Crisis, Stanford, CA: Hoover Institution Press.
Taylor, John, and John Williams, 2009, A black swan in the money market, American Economic Journal—Macroeconomics 1, 58–83.
Tchana, Fulbert, 2007, The welfare cost of banking regulation, Working Paper, University of Montreal, October.
Thakor, Anjan. 1996. Capital requirements, monetary policy, and aggregate bank lending: Theory and empirical evidence. Journal of Finance 51, 279–324.
Tieman, Alexander, 2004, Interest rate pass-through in Romania and other central European economies, IMF Worker Paper WP/04/211.
Tirtiroğlu, Doğan, Kenneth Daniels, and Ercan Tirtiroğlu, 2005, Deregulation, intensity of competition, industry evolution, and the productivity growth of U.S. commercial banks, Journal of Money, Credit, and Banking 37, 339–360.
Toolsema, Linda, Jan-Egbert Sturm, and Jakob de Haan, 2001, Convergence of monetary transmission in EMU: New evidence, CESifo Working paper No. 465.
Van den Heuvel, Skander, 2008, The welfare cost of bank capital requirements, Journal of Monetary Economics 55, 298–320.
VanHoose, David, 2009, Two-sided bank card payment networks and public policy, Lydian Payments Journal, 1, Forthcoming.
VanHoose, David, 2008, Bank capital regulation, economic stability, and monetary policy: What does the academic literature tell us? Atlantic Economic Journal 36, 1–14.
VanHoose, David, 2007a, Market discipline and supervisory discretion in banking: Reinforcing or conflicting pillars of Basel II? Journal of Applied Finance 17, 105–118.
VanHoose, David, 2007b, Theories of bank behavior under capital regulation, Journal of Banking and Finance 31, 3680–3697.
VanHoose, David, 2000, Central bank policy making in competing payment systems, Atlantic Economic Journal 28, 117–139.
VanHoose, David, 1988a, Deregulation and oligopolistic rivalry in bank deposit markets, Journal of Banking and Finance 12, 379–388.
VanHoose, David, 1988b, Deposit market deregulation, implicit deposit rates, and monetary policy, Atlantic Economic Journal 16, 11–23.
VanHoose, David, 1985, Bank market structure and monetary control, Journal of Money, Credit, and Banking 17, 298–311.
VanHoose, David, 1984, Bank market structure and monetary policy, Unpublished Ph.D. Dissertation, University of North Carolina at Chapel Hill.
van Leuvensteijn, Michiel, Christoffer Kik Sørensen, Jacob Bikker, and Adrian van Rixtel, 2008, Impact of bank competition on the interest rate pass-through in the euro area, European Central Bank Working Paper No. 885, March.
Van Roy, Patrick, 2005, The impact of the 1988 Basel accord on banks' capital ratios and credit risk-taking: An international study, Manuscript, European Centre for Advanced Research, University of Brussels, Belgium, July.
Verrecchia, Robert, 1983, Discretionary disclosure, Journal of Accounting and Economics 5, 179–194.
Viscusi, W. Kip, Joseph Harrington, and John Vernon, 2005, Economics of Regulation and Antitrust, Cambridge, MA: MIT Press.
Vives, Xavier, 2008, Competition and regulation in banking: Overview, in Anjan Thakor and Arnoud Boot, eds., Handbook of Financial Intermediation and Banking, Amsterdam: Elsevier, pp. 441–447.
Vives, Xavier, 2001, Restructuring financial regulation in the European Monetary Union, Journal of Financial Services Research 19, 57–82.

Von Rheinbaben, Joachim, and Martin Ruckes, 2004, The number and closeness of bank relationships, Journal of Banking and Finance 28, 1597–1615.
Wagster, John, 1999, The Basel accord of 1988 and the international credit crunch of 1989–1992, Journal of Financial Services Research 15, 123–143.
Walker, Michael, 2005, Merger simulation models: Useful or just plain dangerous, Presentation Notes, Global Competition Law Centre Seminar, College of Europe, October 21.
Wall, Larry, and Robert Eisenbeis, 2000, Financial regulatory structure and the resolution of conflicting goals, Journal of Financial Services Research 17, 223–245.
Wall, Larry, and David Peterson, 1987, The effect of capital adequacy guidelines on large bank holding companies, Journal of Banking and Finance 11, 581–600.
Wallace, Neil, 1989, Another attempt to explain an illiquid banking system: The Diamond-Dybvig model with sequential service taken seriously, Federal Reserve Bank of Minneapolis Quarterly Review 12, 3–16.
Walter, Ingo, 2004, Mergers and Acquisitions in Banking and Finance, Oxford: Oxford University Press.
Walter, John, 2004, Closing troubled banks: How the process works, Federal Reserve Bank of Richmond Economic Quarterly 90, 51–67.
Walter, John, and Patricia Wescott, 2008, Antitrust analysis in banking: Goals, methods, and justifications in a changed environment, Federal Reserve Bank of Richmond Economic Quarterly 94, 45–72.
Wang, Lang, 2005, Bank capital requirements and the effectiveness of monetary policy. Manuscript, Peking University.
Weill, Laurent, 2008, Convergence in banking efficiency across European countries, Manuscript, University of Robert Schuman, January.
Weinberg, John, 2002, Competition among bank regulators, Federal Reserve Bank of Richmond Quarterly 88 (Fall), 19–36.
Whalen, Gary, 2008, Why do de novo banks choose a national charter? Working Paper, Office of the Comptroller of the Currency, December.
Whalen, Gary, 2002, Charter flips by national banks, Office of the Comptroller of the Currency Economic and Policy Analysis Working Paper 2002-1, June.
Wheelock, David, and Paul Wilson, 2000, Why do banks disappear? The determinants of U.S. bank failures and acquisitions, Review of Economics and Statistics 82, 127–138.
Wheelock, David, and Paul Wilson, 1999, Technical progress, inefficiency, and productivity change in U.S. banking, 1984–1993, Journal of Money Credit, and Banking 31, 212–234.
Whinston, Michael, 2007, Antitrust policy toward horizontal mergers, in Mark Armstrong and Robert H. Porter, eds., Handbook of Industrial Organization, Volume 3, Amsterdam: Elsevier North Holland, pp. 2371–2440.
Whinston, Michael, 2006, Lectures on Antitrust Economics, Cambridge, MA: MIT Press.
White, Lawrence, 2006, Market definition and market power in payment card networks: Some comments and considerations, Review of Network Economics 5, 61–71.
White, Lawrence, 1991, The S&L Debacle: Public Policy Lessons for Bank and Thrift Regulation, Oxford: Oxford University Press.
White, Lawrence, 1976, Price regulation and quality rivalry in a profit-maximizing model: The case of bank branching, Journal of Money, Credit, and Banking 8, 97–106.
Wicker, Elmus, 1996, The Banking Panics of the Great Depression, Cambridge: Cambridge University Press.
Winker, Peter, 1999, Sluggish adjustment of interest rates and credit rationing: An application of unit root testing and error correction modeling, Applied Economics 31, 267–277.
Winston, Clifford, 2006, Government Failure versus Market Failure, Washington, DC: AEI Brookings.
Wolken, John, and Frederick Derrick, 1986, Advertising, market power, and non-price competition: Evidence from commercial banking, Board of Governors of the Federal Reserve System, Working Paper No. 86-3.

Wood, John, 1975, Commercial Bank Loan and Investment Behavior, London: John Wiley and Sons.
Wright, Julian, 2004, The determinants of optimal interchange fees in payment systems, Journal of Industrial Economics 52, 1–26.
Wright, Julian, 2003, Optimal card payment systems, European Economic Review 47, 587–612.
Yafeh, Yishay, and Oved Yosha, 2001, Industrial organization of financial systems and strategic use of relationship banking, European Finance Review 5, 63–78.
Zarutskie, Rebecca, 2006, Evidence on the effects of bank competition on firm borrowing and investment, Journal of Financial Economics 81, 503–537.

Index