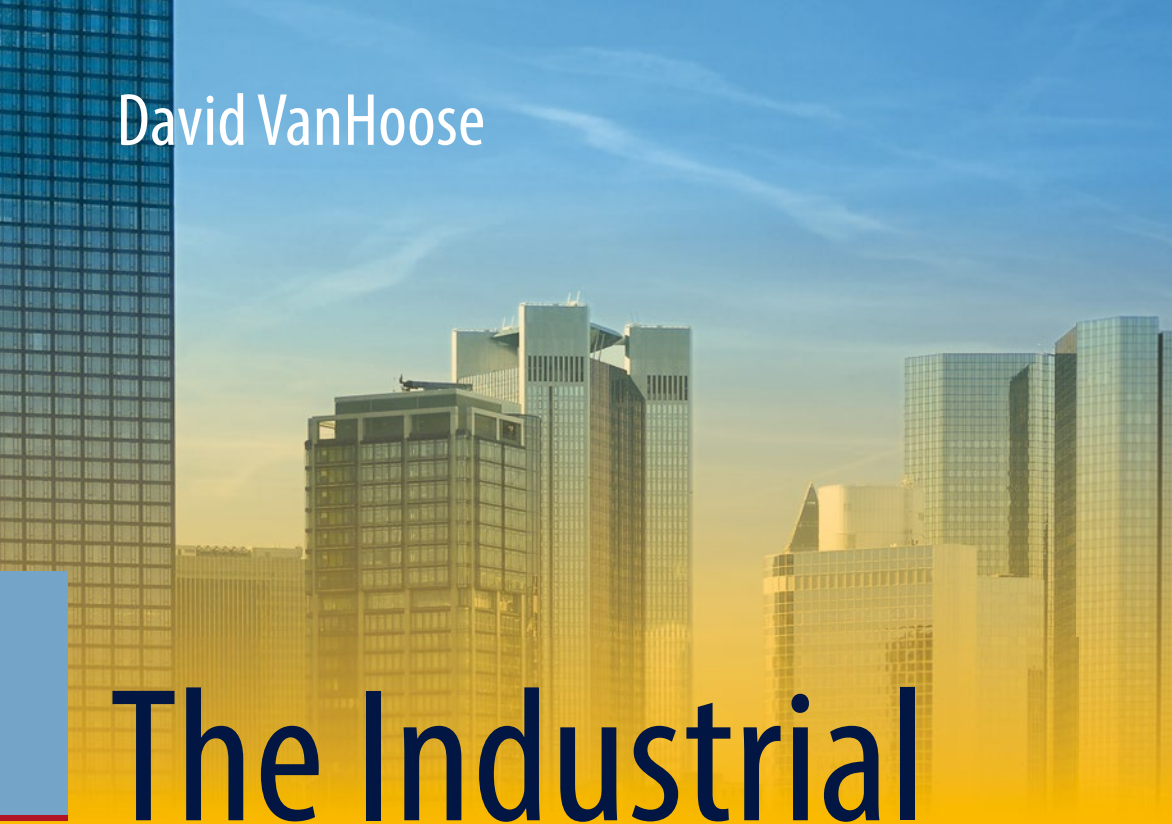

David VanHoose

The Industrial Organization of Banking

Bank Behavior, Market Structure, and Regulation

Second Edition

Springer

The Industrial Organization of Banking

David VanHoose

The Industrial Organization of Banking

Bank Behavior, Market Structure, and Regulation

Second Edition

Springer

David VanHoose
Waco, Texas
USA

ISBN 978-3-662-54325-2 ISBN 978-3-662-54326-9 (eBook)
DOI 10.1007/978-3-662-54326-9

Library of Congress Control Number: 2017935017

© Springer-Verlag GmbH Germany 2010, 2017
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer-Verlag GmbH Germany
The registered company address is: Heidelberger Platz 3, 14197 Berlin, Germany

Contents

1 Introduction: The Banking Environment 1
Three Fundamental Areas Within the Industrial Organization of Banking 1
Objectives 2
Bank Behavior and the Structure of Banking Markets 2
Bank Competition and Public Policy 3
Assessing Bank Regulation 3
The Bank Balance Sheet 4
Bank Assets 4
Loans 5
Securities 6
Cash Assets 6
Trends in U.S. Bank Asset Allocations 6
Bank Liabilities and Equity Capital 8
Large-Denomination Time Deposits 8
Transactions Deposits, Savings Deposits, and Small-Denomination Time Deposits 8
Purchased Funds and Subordinated Notes and Debentures 9
Bank Capital 9
Trends in Bank Liabilities and Equity Capital 9
The Bank Income Statement 10
Interest Income 11
Noninterest Income 11
Interest Expenses 11
Expenses for Loan Loss Provisions 12
Real Resource Expenses 12
Bank Profitability Measures 12
Asymmetric Information and Risks in Banking 13
Adverse Selection 14
Moral Hazard 14

Risks on the Balance Sheet 14
Credit Risk 15
Market Risks 15
Liquidity Risk 15
Systemic Risk 16
Trends in U.S. Banking Industry Structure 16
The Number of Commercial Banks 16
Mergers, Acquisitions, and Concentration 16
References 18

2 Alternative Perspectives on Bank Behavior 21
Identifying the Outputs and Inputs of a Bank 21
What Banks Do: Alternative Perspectives on Bank Production 21
Assessing the Economic Outputs and Inputs of Banks 22
Banks as Portfolio Managers 24
The Basic Bank Portfolio-Management Model 24
Limitations of Portfolio Management Models 25
Banks as Firms 25
A Perfectly Competitive Banking Industry 26
A Static Banking Model 26
Fundamental Dynamics in a Perfectly Competitive Banking Model and Implications for Portfolio Separation 32
Imperfectly Competitive Banking Markets 35
Monopolistic and Monopsonistic Interest Rate Determination in Bank Loan and Deposit Markets 35
Social Losses Due to Imperfect Competition in Banking 37
Alternative Modes of Behavior Between Perfect Competition and Monopoly and Monopsony 38
References 43

3 The Industrial Economics of Banking 47
The Structure-Conduct-Performance Paradigm in Banking 47
The SCP Hypothesis with Identical Banks 48
Structural Asymmetry, Dominant Banks, and the SCP Paradigm . . . 49
A Dominant-Bank Model 49
Strategic Entry Deterrence 51
Evaluating the Applicability of the SCP Paradigm to the Banking Industry 51
Traditional SCP Evidence from Cross-Sectional Banking Data . . . 52
Evidence from Cross-Country Studies 53
Dynamic Interest Rate Responses: Competition and Pass-Through Effects 53
The Conduct and Relative Performances of Large and Small Banks 55

Market Structure and Bank-Customer Relationships 57
Basic Market-Structure Implications of Bank-Customer Relationships 58
Evidence on Bank-Customer Relationships 60
Determinants and Impacts of Bank-Customer Relationships 60
Competition and Relationship Lending 63
The Efficient Structure Theory and Banking Costs 65
The Efficient Structure Challenge to the SCP Paradigm 65
Banking Efficiency and Costs 66
Efficient Structure Theory and Bank Performance 68
Endogenous Sunk Fixed Costs and Banking Industry Structure 71
Endogenous Sunk Costs and Concentration 72
Non-Price Competition in Banking: Implicit Deposit Rates Versus Quality Rivalry 73
Evidence on Advertising Outlays in the Banking Industry 75
Endogenous Sunk Costs and the Banking Industry 76
References 77

4 The Economics of International Banking 89
The Growth of Global Banking 90
The First Big Wave of Banking Globalization: The Colonial Period 90
The Second Globalization Wave: The Most Recent Decades 90
Fundamental Elements Influencing Cross-Border Banking 92
Distance 92
Competitive Structure and Effects of Entry in Host-Country Banking Markets 93
Bank Size and Entry Costs 95
Idiosyncratic Elements and Prior Experience 96
The Industrial Organization of Multinational Banking 97
Determinants of the Overall Scope of Foreign Direct Investment in Banking 97
Greenfield Investment Versus Acquisitions of Host Nations' Banks . . . 98
Real-Resource-Based Analyses of Cross-Border Trade in Banking Services 100
Differences in Nation's Capital-Labor Ratios and Efficiencies as Determinants of Cross-Border Banking 102
The Structure of Niepmann's Framework for a Closed Economy . . . 102
Extending Niepmann's Setup to a Two-Country Environment . . . 103
An Intra-industry Model of International Trade in Banking Services 105
A Closed-Economy Setting with Differentiated Banking Services 106
A Two-Economy Environment with Traded Banking Services . . . 107
Evidence on the Empirical Relevance of Trade-Based Theories of International Banking 109
References 110

5 The Economics of Banking Antitrust 115
Why Banks Merge 115
Profit Enhancements from Mergers 115
Diversification Benefits of Bank Mergers 118
Assessing Loan and Deposit Market Effects of Bank Consolidation . . . 120
Mergers in Initially Perfectly Competitive Banking Markets 120
Mergers in Initially Imperfectly Competitive Banking Markets 122
Evidence on the Consequences of Banking Consolidation 123
Mergers and Market Power 123
Evidence on Efficiency Gains from Banking Consolidation 126
Banking Antitrust in Practice 127
U.S. Bank Merger Guidelines 127
The Relevant Market 128
Merger Screening 129
Evaluating the U.S. Bank Merger Guidelines 130
Is the Official Relevant Banking Market Really Relevant? 131
Do the Formal Guidelines Mis-Measure Market Power? 133
Implications of Endogenous Sunk Fixed Costs 134
Do Banking Consolidations Preclude Entry and Reduce Consumer Welfare? 135
Rethinking Bank Merger Analysis 136
References 138

6 Bank Competition, Stability, and Regulation 143
Banks as Issuers of Demandable Debt 144
The Diamond-Dybvig Model 144
An Optimal Risk-Sharing Contract 145
The Diamond-Dybvig Intermediation Solution and the Problem of Runs 145
Evaluating the Diamond-Dybvig Analysis 146
Banks as Screeners and Monitors 148
Evidence on Bank Monitoring Activities 149
Evidence from Announcement Effects 150
Evidence from Firm Investment and Bond Yields 151
Evidence from Syndicated Loans and Loan Sales 151
Direct Evidence of Bank Monitoring Activities 152
A Monitoring Model with Heterogeneous Banks 153
Behavior of Monitoring and Nonmonitoring Banks 153
Loan Market Equilibrium and Equilibrium Monitoring 154
The Relationship Between Banking Competition and Risks 156
Perfect Competition and Bank Risks 156
The "Excessive Deposit Competition" Argument 157
The Competition-Illiquidity Argument 158
The Competition-Asset Risk Argument 159

Market Power and Bank Risks: Theory and Evidence 160
Competition and Risk: Theory 161
Bank Size, Competition, and Risk: Evidence 162
Deposit Insurance, "Too Big to Fail" Doctrine, and the Basel Standards 164
The Distorting Effects of Deposit Insurance 164
The Impact of the Too-Big-to-Fail Doctrine 166
Capital Regulation and the Three Pillars of the Basel Framework 169
References 169

7 Capital Regulation, Bank Behavior, and Market Structure 175
The Portfolio Management Perspective on Capital Regulation 176
The Bank as a Competitive, Mean-Variance Portfolio Manager Facing Capital-Constrained Asset Portfolios 176
Taking Deposit Insurance Distortions into Account 178
Explaining the Mixed Implications of Portfolio Management Models 180
Asset-Liability Management Under Capital Regulation 181
An Incentive-Based Perspective on Capital Regulation 182
Incentives and Capital Requirements 182
Perfect Competition Models of Bank Capital Regulation 183
Monopolistic Competition Models of Capital Regulation 187
Demandable Debt, Bank Risks, and Capital Regulation 189
Capital Regulation and Fragile Deposits 189
Moral Hazard, Bank Lending and Monitoring, and Capital Regulation 191
Capital Regulation and Bank Heterogeneities 192
Adverse Selection and Capital Regulation 193
Capital Requirements, Heterogeneous Banks, and Industry Structure 194
Capital, Shocks and Procyclicality, and Bank Performance 198
Does Toughening Capital Requirements Boost Bank Capital Ratios and Create Credit Shocks? 199
Procyclical Features of a Capital-Regulated Banking Industry 201
Empirical Evidence on Procyclical Effects of Capital Regulation . . . 203
Do Tougher Capital Regulations Yield Better Outcomes? Empirical Evidence 204
References 206

8 Market Discipline and the Banking Industry 213
The Basel Framework's Market Discipline Pillar 214
The Channels of Market Discipline 215
Motivations of Agents Who Discipline Banks 215
Conditions for Market Signals to Effectively Discipline Banks 217

Potential Benefits and Costs of Market Discipline in Banking 218
Evaluating Incentives for Information Disclosure 218
Ways to Enhance Bank Market Discipline 220
Industry Structure and Market Discipline 223
Market Discipline in a Basic Banking Model 223
Market Power, Information Disclosure, and Market Discipline 224
Evidence on Market Discipline's Effectiveness 227
Information Content of Market Prices and Bond Yield Spreads 227
Market Discipline Versus Regulation 229
Regulatory Crowding Out of Market Discipline? 230
Additional Evidence on Interactions between Regulation and Market Discipline 231
Evidence on Bank Information Disclosure 233
Evaluating the Market Discipline Pillar vis-à-vis the Other Pillars of the Basel Framework 234
The Limitations of Market Discipline under the Basel Framework . . . 234
Theory Versus Reality Under the Basel Market Discipline Pillar . . . 235
References 236

9 Regulation and the Structure of the Banking Industry 243
Public Interest Versus Public Choice Perspectives on Bank Regulation 243
Public Interest and the Alleged "Need" for Bank Regulation 243
Public Choice Motivations for Bank Regulation 244
Applying the Economic Theory of Regulation to the Banking Industry 245
Assessing the Implications of the Economic Theory of Regulation 246
A Generalized Perspective on Evaluating Bank Regulation 249
The Political Economy of Banking Supervision Conducted by Multiple Regulators: Is a "Race to the Bottom" Unavoidable? 250
Regulatory Preferences and Bank Closure Policies 250
Competition Among Bank Regulators 252
A Theory of Optimal Supervisory Choices of a Single Bank Regulator 253
The Case of Competing Regulators 255
A Supervisory Race to the Bottom? 255
Is Greater Centralization of Regulatory Functions the Answer? . . . 257
The Supervisory Review Process Pillar of the Basel Standards 259
The Supervisory Review Process Pillar: Conceptual Issues 261
Discretion Versus Rules 261
How Tough Should a Supervisory Policy Rule Really Be? 262
Is There Really a Basel Supervisory Review Process? 264
Regulatory Compliance Costs and Industry Structure 266

Assessing Banks' Costs of Basel Compliance: Economies of Regulation? 266
Rule-of-Thumb Estimates 266
Estimates of Basel Compliance Costs Based on Survey Data 267
"Economies-of-Basel"—Scale Advantages in Basel Compliance? 268
To What Extent Do Basel Compliance Costs "Matter"? 269
Bank Regulation and Endogenous Fixed Costs 270
Regulatory Compliance Costs: A Missing Component? 271
Regulatory Sunk Fixed Costs 272
References 275

10 Macroprudential Regulation and International Policy Coordination 281
Systemic Risk 281
Defining Systemic Risk 281
Measuring Systemic Risk 283
Doing Something About Systemic Risk: Macroprudential Policy 284
Implementation of Macroprudential Policy 284
Structuring Macroprudential Regulation 285
Additional Macroprudential Policy Instruments 286
Pitfalls of Macroprudential Regulation 288
International Regulatory Policy Coordination 289
Essentials of the Regulatory Policy Coordination Problem 289
Recent Research on International Coordination of Financial Regulatory Policies 291
Can Impediments to International Bank Regulatory Coordination Be Surmounted? 296
References 299

Chapter 1
Introduction: The Banking Environment

Stocks, Flows, Information, and Risks

This second edition of The Industrial Organization explores the study of the structure of individual banks, banking markets, and their interactions. The book has the same two key objectives as pursued in the first edition. The first goal is to assist students and policymakers in climbing the field's steep learning curve as effectively as possible. The second is to provide a full survey of the field 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, assesses current understanding of the relationship between bank market structure and the stability of the banking industry, and contemplates determinants of international banking activities. Finally, it evaluates the issue of systemic risks in banking, considers the implications of bank capital regulation, appraises the potential interaction between market discipline and direct regulatory supervision of banks, explores the interplay between regulation and the structure of the banking industry, and assesses issues regarding macroprudential regulation and the potential international coordination of bank regulatory and supervisory policies.

Three Fundamental Areas Within the Industrial Organization of Banking

The book focuses on three fundamental areas of study within 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; and

© Springer-Verlag GmbH Germany 2017

D. VanHoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-662-54326-9_1

(3) assessing the implications of proposed or actual regulations for individual banks and/or the banking industry.

Researchers typically contemplate issues relating to one or perhaps two of these areas but only rarely all three. It can prove difficult, therefore, for a student or a policymaker seeking to learn about the industrial organization of banking to locate a single source regarding 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 (2015), Degryse et al. (2009), and Matthews and Thompson (2013) or survey articles covering specific topic areas that are scattered across a handful of issues of journals and books containing collected readings.

Objectives

This book's fundamental purpose is to provide a complete overview of the economic profession's current understanding of the interplay among bank behavior, market structure, and regulation. One key aim 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 is to synthesize diverse strands of the banking literature at a level appropriate for bankers and policymakers.

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 framework for banking supervision, because most bank regulators have maintained a steadfast devotion to the principles entailed in this framework, even in the wake of the events encompassing and following the panic of the late 2000s.

Bank Behavior and the Structure of Banking Markets

The present chapter reviews 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 are discussed later in the present chapter. The chapter also surveys recent trends in the structure of banking revealed by data from U.S. commercial banks.

Chapter 2 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. It considers perfectly competitive behavior, monopoly in

loan markets and monopsony in deposit markets, standard Cournot-Nash and Bertrand-Nash models of bank behavior in oligopolistic settings with homogenous loans and deposits, and monopolistic competition with differentiated loans and deposits.

Chapter 3 applies the theories introduced in Chap. 2 to discussion of and evaluation of alternative approaches to the industrial economics of banking. It shows how the static imperfect-competition frameworks discussed in Chap. 2 can provide a foundation for the structure-conduct-performance (SCP) and efficient-structure paradigms. It also reviews recent work on endogenous sunk fixed costs in banking.

Chapter 4, which is new to this edition, examines banking in an international context. The chapter begins by reviewing experience with and surveying fundamental elements influencing cross-border banking. Next, it evaluates traditional perspectives on determinants of foreign direct investment on the part of banks. The chapter concludes by discussing recent research analyzing cross-border trade in banking services from the point of view of real-resource-based models of international trade.

Bank Competition and Public Policy

Chapter 5 considers 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.

Chapter 6 focuses on the implications of market structure and competition for stability of the banking industry. It presents and evaluates theories of banks as issuers of demandable debt and considers their how loan monitoring activities can be incorporated into basic banking theory and reviews evidence regarding the empirical importance of monitoring activities. It concludes by discussing aspects of active governmental involvement in the banking industry intended to improve its stability prospects.

Assessing Bank Regulation

Chapter 7 focuses on how industrial organization shapes the impacts of bank capital regulation formalized under the Basel 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.

Chapter 8 considers the role of market discipline in the banking industry. The chapter begins by providing a basic overview of the Basel 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 closes with an evaluation of the Basel market discipline pillar in relation to the capital standards and supervisory process pillars.

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, surveys research on optimal bank closure policies, and considers competition among government regulators facing overlapping jurisdictions of clienteles that can choose which of the regulators serve as their primary supervisor. 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.

Chapter 10, which also is new to this edition, concludes by contemplating recent issues regarding macroprudential regulation aimed at containing systemic risk and the related issue of international policy coordination. After discussing the definition and measurement of systemic risk, it considers proposals for and potential pitfalls of macroprudential regulation. The remainder of the chapter contemplates potential gains and losses from international coordination of banking policies, hindrances to effective coordination, and possible ways to surmount such impediments.

The Bank Balance Sheet

The 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 is an obligation by another party to repay principal plus any contracted interest to the bank within a specified period. Table 1.1 lists the combined assets of all domestically chartered U.S. commercial banks.

Table 1.1 Assets of U.S. commercial banks

Asset category	$ Billions	%
Commercial and industrial loans	2064.8	12.9
Consumer loans	1349.7	8.4
Real estate loans	4075.3	25.4
Interbank loans	62.9	0.4
Other loans	*1508.0*	*9.4*
Total loans	9060.7	56.5
Securities	3290.1	20.5
Cash assets	2354.4	14.7
Other assets	*1318.6*	*8.3*
Total assets	16,023.8	100.0

Source: Board of Governors of the Federal Reserve System, September 2016

Loans

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 extended to business enterprises account for more than 12% of total bank assets.
- ***Consumer Loans*** Table 1.1 shows that consumer loans account for 9% of U.S. bank assets. 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 1- to 5-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% in 1985 to more than 60% in the mid-2000s before dropping closer to previous historical norms.
- ***Interbank Loans*** Banks lend funds to each other directly in interbank loan markets, such as the U.S. federal funds market in which banks borrow from and lend to each other deposits held at Federal Reserve banks. Banks typically extend interbank loans in large-denomination units ranging from $200,000 to well over $1 million per loan. U.S. interbank lending plummeted after the Federal Reserve implemented in October 2008 a policy of paying interest on both required and excess reserve balances higher than the prevailing U.S. interbank (federal funds) rate (see Dutkowsky and VanHoose 2017).

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

U.S. government securities, including Treasury bills, notes, and bonds, account for just over 20% 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.

Correspondent balances, or funds that banks hold on deposit with other private, correspondent banking institutions, are the second type of cash asset. The third 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.

The fourth and most important form of cash asset is reserves held with the central bank, such as reserve deposits that U.S. banks maintain with Federal Reserve banks. Banks transmit 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. A massive expansion of excess reserve holdings since the Federal Reserve's October 2008 implementation of its interest-on-reserves policy generated a roughly tenfold boost in the percentage of bank assets held as cash assets.

Trends in U.S. Bank Asset Allocations

Figure 1.1 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.

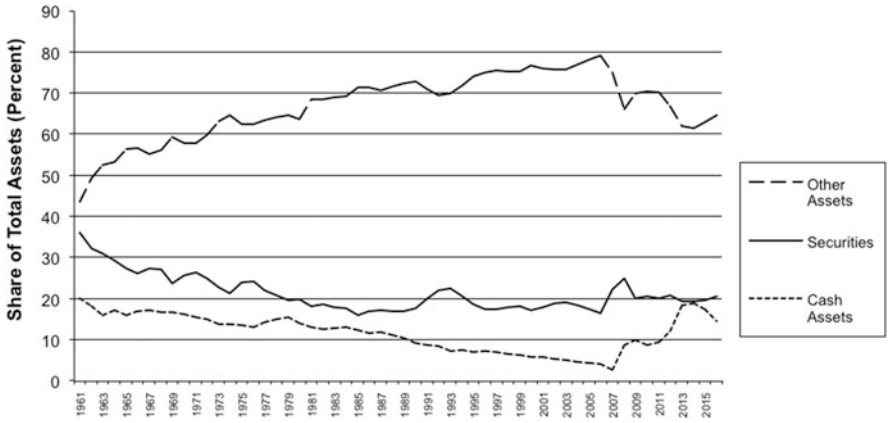

Fig. 1.1 U.S. commercial banks' asset allocations (Source: Board of Governors of the Federal Reserve System)

Allocations to assets other than securities and cash assets—primarily loans—rose markedly into the late 1980s.

There was a general downward trend in relative holdings of cash assets until 2007. As noted above, the Federal Reserve's payment of interest on all reserves since October 2008 generated an upsurge in holdings of cash assets. Bank security holdings as a share of total assets also exhibited a slight downward trend through the later 1980s before stabilizing at about 20%.

Figure 1.2 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%. 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 through the late 2000s, U.S. banks' focus shifted to real estate lending. During the years since the 2007–2010 housing meltdown, this focus has diminished.

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

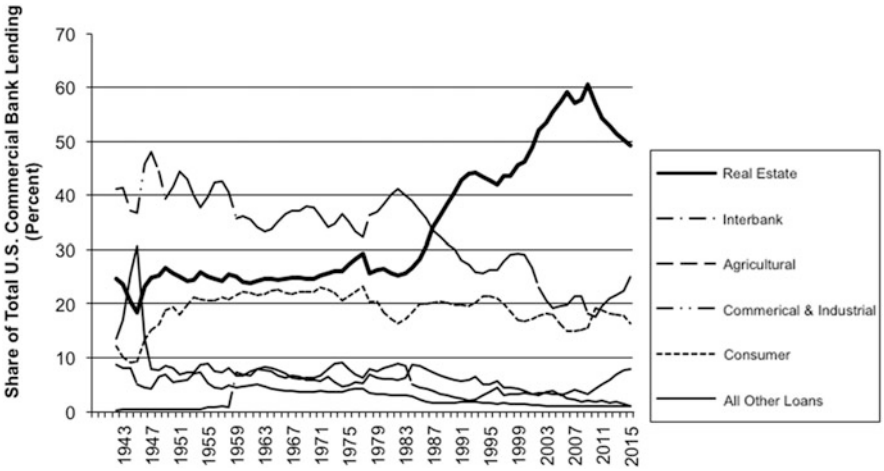

Fig. 1.2 U.S. commercial banks' loan allocations (Source: Federal Deposit Insurance Corporation)

Bank Liabilities and Equity Capital

A liability of a bank is the value of a legal claim on its assets. Table 1.2 lists the combined total liabilities and equity capital all U.S. banks.

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 have 6-month terms and trade actively. Large CDs and other large-denomination time deposits account for just over 9% of bank liabilities and equity capital.

Transactions Deposits, Savings Deposits, and Small-Denomination Time Deposits

Transaction deposit accounts are accounts from which owners may draw funds via checks or debit cards. 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. Taken together, these deposits are the "other" deposits listed in Table 1.2 as comprising more than 61% of banks' total liabilities and equity capital.

Table 1.2 U.S. commercial banks' liabilities and equity capital

Category	$ Billions	%
Large time deposits	1509.6	9.4
Other deposits	*9847.2*	*61.5*
Total deposits	11,356.8	70.9
Borrowings	2005.7	12.5
Other liabilities	*948.3*	*5.9*
Total liabilities	14,310.8	89.3
Equity capital	1713.0	10.7
Total liabilities and equity capital	16,023.8	100.0

Source: Board of Governors of the Federal Reserve System, September 2016

Purchased Funds and Subordinated Notes and Debentures

Key liabilities among the "borrowings" and "other liabilities" categories in Table 1.2 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 debt instruments with maturities in excess of one year. Those who hold these debt instruments have subordinated claims in the event of bank failures. In the event of bankruptcy, holders of subordinated notes and debentures would receive no payments until all depositors 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 Chaps. 6 and 7, 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 risen above 10%.

Trends in Bank Liabilities and Equity Capital

Figure 1.3 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 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 as high as nearly 30% of total liabilities and equity capital.

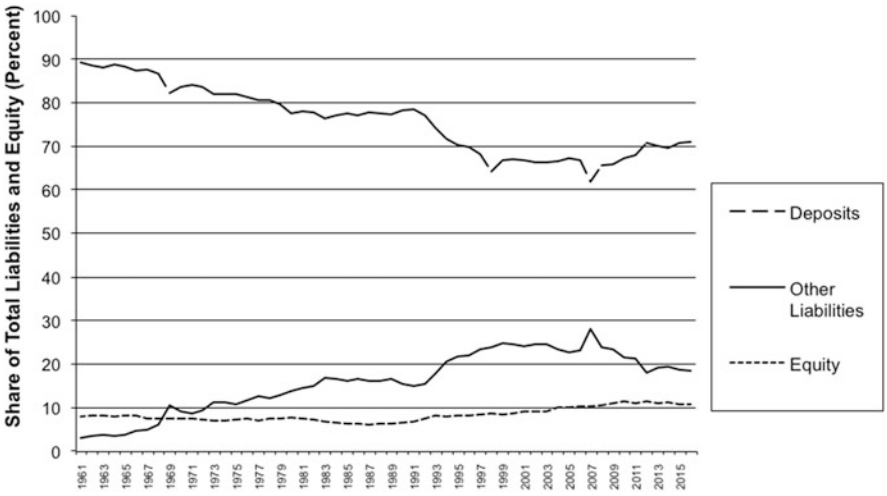

Fig. 1.3 U.S. commercial banks' liabilities and equity capital (Source: Board of Governors of the Federal Reserve System)

A key reason for the shift from deposits to purchased funds was that banks struggled to attract sufficient deposits to fund desired their desired asset scales. Savers could earn higher yields by holding other instruments such as government securities, so banks borrowed from other sources to fund some of their operations. Raising equity funds 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 Chap. 7.

What difference does it make what source of funds banks utilize? Based on data from more than 1300 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 incomes, or revenues, as flows over time. Hence, they tabulate and report interest income in quarterly and annual income statements.

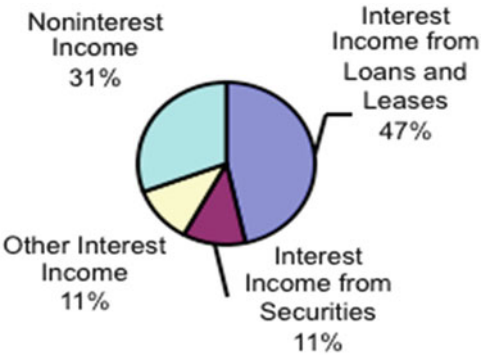

Fig. 1.4 Sources of U.S. commercial banks' revenues (Source: Federal Deposit Insurance Corporation)

Interest Income

Figure 1.4 shows that interest income accounts for about two-thirds of U.S. commercial banks' revenues. The bulk of interest income is derived from loan interest income, which accounts for just over half of total earnings.

Noninterest Income

As Fig. 1.4 indicates, U.S. commercial banks earn slightly more than 40% of their revenues as noninterest income, such as trading profits, customer service charges, and loan management fees.

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. 1.5, interest expenses account only 9% of the total costs incurred by U.S. commercial banks. This amount is about 30% points lower as compared with the late 2000s, prior to the recent decrease in market interest rates.

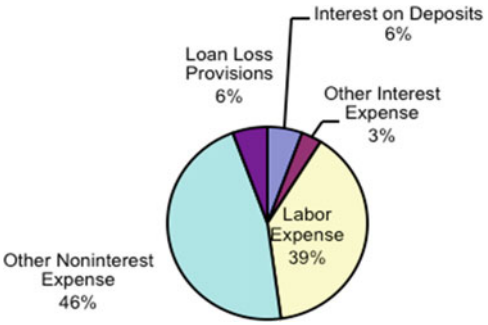

Fig. 1.5 U.S. commercial banks' expenses (Source: Federal Deposit Insurance Corporation)

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. Loan loss provisions have recently accounted for about 6% of U.S. banks' expenses.

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.

Expenses on real resources amount to more than 80% of costs incurred by U.S. commercial banks. Clearly, real resource expenditures are a substantial portion of banks' total costs. Even prior to the low-interest-rate environment experienced recently, spending on real resources accounted for at least one-half of banks' expenditures.

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.

Figure 1.6 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. The net interest margin recovered following the financial meltdown and associated economic slump.

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 several 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 aggregate 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.

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 those on 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 earnings volatility, 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.

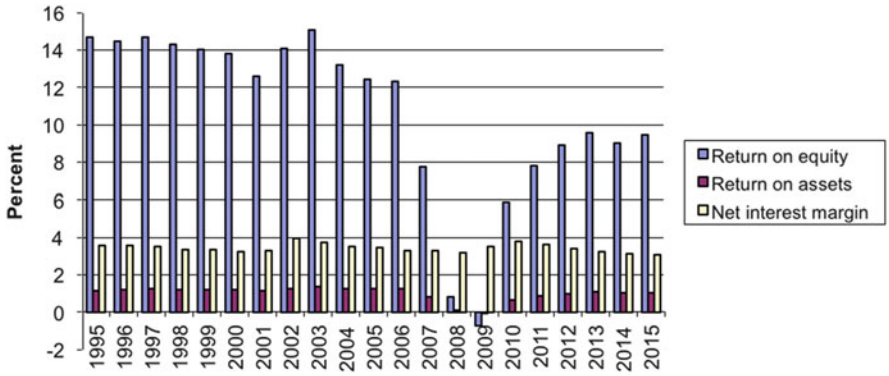

Fig. 1.6 U.S. commercial banks' average return on assets and return on equity (Source: Federal Deposit Insurance Corporation)

Adverse Selection

Adverse selection is the potential for those who desire funds for undeserving projects to be among 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

Moral hazard arises because after credit has been obtained 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, a several 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. 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

Banks hold a variety of securities alongside their loans, and they encounter market risks on both types of assets. One manifestation of market risk is exposure to price risk, or the potential for a 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 from the potential for interest rates on liabilities to rise more rapidly than 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. For 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 are contemplated in Chaps. 5 and 10.

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 de posits; and on bank profits and risks. As discussed in Chaps. 3 and 4, changes in banking structures have enabled researchers to explore these effects in considerable detail.

The Number of Commercial Banks

Figure 1.7 shows that since the mid-1980s, the number of U.S. commercial banks has dropped by about 70%. 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% of the combined assets of all commercial banks. Today, fewer than 30% of banks have total assets below $100 million, and their combined assets make up less than 1% of the consolidated assets of the industry.

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. The primary explanation for the decline in the number of U.S. banks since the early 1980s has been mergers, not bank closures, as can be discerned from Fig. 1.8.

Each year from 1990 through 2001, mergers and acquisitions redistributed more than 2% of aggregate bank assets in the United States. Hundreds of billions of dollars of assets have been reallocated via merger every year. Thus, even though

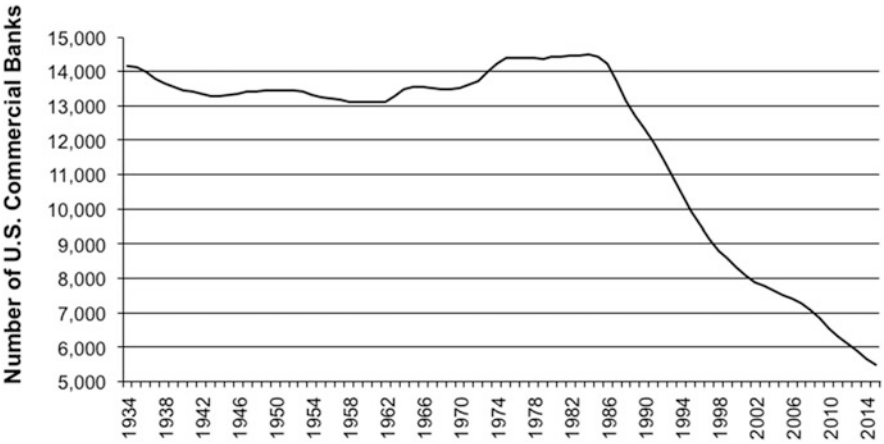

Fig. 1.7 The Number of U.S. Commercial Banks Since 1934 (Source: Federal Deposit Insurance Corporation)

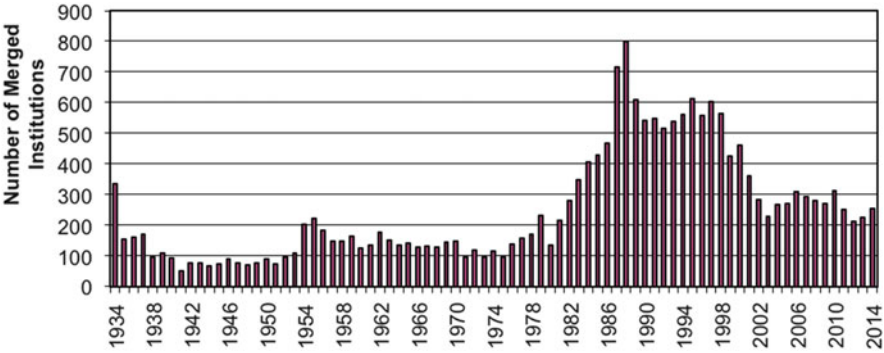

Fig. 1.8 Number of U.S. bank mergers (Source: Federal Deposit Insurance Corporation)

much of this merger-and-acquisition activity has involved larger banks gobbling up smaller institutions, a considerable portion also involved combinations of larger banks.

A consequence of this bank merger-and-acquisition wave has been a rise in aggregate industry concentration. As shown in Fig. 1.9, the percentages of deposits held at the largest U.S. banks generally has increased since 1990. This reflects a trend, documented by Janicki and Prescott (2006), Jones and Critchfield (2008), and Kowalik et al. (2015), of a shift in the size distribution of U.S. banks toward larger banking organizations.

How do economists take into account the relatively more concentrated nature of banking markets when studying the behavior of individual banks? What are economic implications of the trend toward larger banking institutions and greater market concentration? Are more concentrated and potentially less competitive

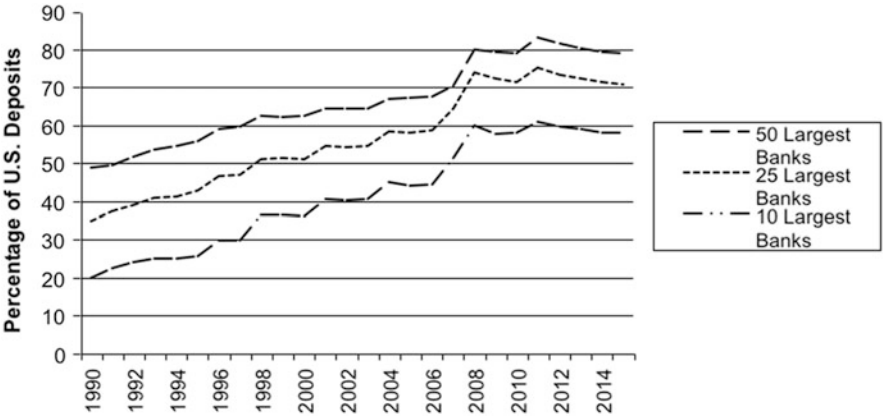

Fig. 1.9 Nationwide deposit concentration in the U.S. banking industry (Sources: Pilloff (2009) and Board of Governors of the Federal Reserve System)

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 be more stable? These are key questions explored in the following chapters.

Acknowledgments I 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 6 incorporates portions of VanHoose (2013), Chap. 7 includes portions of VanHoose (2007b, 2008), portions of Chaps. 8 and 9 draw on VanHoose (2007a), and parts of Chap. 10 build on portions of VanHoose (2011, 2016). 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.

References

Acharya, Viral, Iftekhar Hasan, and Anthony Saunders. 2006. Should banks be diversified? Evidence from individual bank loan portfolios. *Journal of Business* 79: 1355–1412.

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.

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.

Degryse, Hans, Moshe Kim, and Steven Ongena. 2009. *Microeconometrics of Banking: Methods, Applications, and Results*. Oxford: Oxford University Press.

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.

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.

Dutkowsky, Donald, and David VanHoose. 2017. Interest on reserves, regime shifts, and bank lending. *Journal of Economics and Business* 27. (In press).

Freixas, Xavier, and Jean-Charles Rochet. 2008. *Microeconomics of Banking*. 2nd ed. Cambridge, MA: MIT Press.

Greenbaum, Stuart, and Anjan Thakor. 2015. *Contemporary Financial Intermediation*. 2nd ed. Burlington, MA: Academic Press.

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.

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.

Jones, Kenneth, and Tim Critchfield. 2008. Consolidation in the U.S. banking industry: Is the "long, strange trip" about to end? In *Handbook of Financial Intermediation and Banking*, ed. Anjan Thakor and Arnoud Boot, 309–346. Amsterdam: Elsevier.

Kowalik, Michal, Troy Davig, Charles Morris, and Kirsten Regehr. 2015. Bank consolidation and merger activity following the crisis. *Federal Reserve Bank of Kansas City Economic Review*, First Quarter, 31–49

Matthews, Kent, and John Thompson. 2013. *The Economics of Banking*. Chichester, UK: Wiley.

Mercieca, Steve, Klaus Schaeck, and Simon Wolfe. 2007. Small European banks: Benefits from diversification? *Journal of Banking and Finance* 31: 1975–1998.

Pilloff, Steven. 2009. The banking industry. In *The Structure of American Industry*, ed. James Brock, 12th ed. Upper Saddle River, NJ: Pearson.

Stiroh, Kevin. 2004. Diversification in banking: Is noninterest income the answer? *Journal of Money, Credit, and Banking* 36: 853–882.

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. 2008. Bank capital regulation, economic stability, and monetary policy: What does the academic literature tell us? *Atlantic Economic Journal* 36: 1–14.

VanHoose, David. 2011. Systemic risk and macroprudential bank regulation: A critical appraisal. *Journal of Financial Transformation* 33: 45–60.

VanHoose, David. 2013. A model of international trade in banking services. *Open Economies Review* 24: 613–625.

VanHoose, David. 2016. Should financial regulators engage in international policy coordination? *International Economics and Economic Policy* 13: 319–338.

Chapter 2
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 that channels funds from savers to entrepreneurs who make capital investments or to individuals who purchase durable goods or tangible assets. Savers who lend funds to financial intermediaries such as banks otherwise could have chosen to engage in direct finance by lending 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' views begin to diverge. 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 bank behavior. Dewatripont and Tirole (1993) represent a perspective that focuses

© Springer-Verlag GmbH Germany 2017

D. VanHoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-662-54326-9_2

on banks 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 (1) writing and enforcing debt contracts that match savers preferring highly liquid assets with firms desiring to finance capital investment via long-term credits; (2) 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; (3) engaging in delegated screening and monitoring to determine whether prospective borrowers are creditworthy and whether actual borrowers are directing funds to worthwhile projects; and (4) 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 views fit into two approaches to measuring what banks produce. Under one, which Berger and Humphrey (1997) term the production approach and which was first utilized by Benston (1965), banks specialize in producing services for holders of loan and deposit accounts. Hence, the production approach, which receives support from the Dewantripont-Tirole discussion, recommends that measures of a bank's output should focus on numbers of 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 empirically. They note that data on transactions 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 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. As Basu et al. (2011) and Alon et al. (2011) point out, another element that further complicates an assessment of bank outputs is that the fact that banks commonly seek lending compensation through spreads instead of fees makes valuing the banks' output problematic.

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 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, for which Royster (2012) describes a recently updated set of measurement techniques, labor, physical capital, and purchased funds typically are classified as bank inputs.

The third is the user-cost method employed by Hancock (1985, 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—often called "core" deposits, the value of which Sheehan (2013) find varies considerably by institution. 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.

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 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.

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.

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. 2.1. 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 available assets to which banks may allocate deposit funds.

If 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 which 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. Point P corresponds to a specific allocation of various assets as per unit of bank equity. Hence, this point uniquely identifies the asset allocation that maximizes the bank owners' expected utility given their preferences toward expected return and risk.

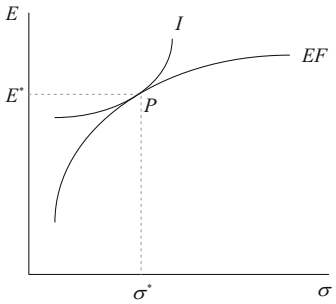

Fig. 2.1 A bank's selection of the optimal portfolio

Limitations of Portfolio Management Models

An 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).

Nevertheless, the assumptions underlying portfolio management models place restrictions on their suitability—at least, absent significant modifications—to industrial organization applications. Banks operate in a variety of markets, including loan and deposit markets 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 industrial organization issues. As noted in the previous chapter, more than half of the costs incurred by U.S. banks are non-interest 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. Portfolio management models also 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 consider a bank that operates within a perfectly competitive industry. This bank is insignificant in size relative to all markets in which it operates and issues liabilities. All traded assets are homogeneous and subject to identical risks. There are no substantial barriers to entry or exit. No informational asymmetries exist, and 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 Chap. 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 $C^i_{ZZ} \equiv \partial^2 C^i / \partial \left(Z^i\right)^2 \geq 0$, 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 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. 2.2 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 bank's marginal return on lending is the market clearing loan rate r_L^*. The bank's total marginal cost of lending, denoted MC_L^i, equals the sum of the interest rate per dollar of deposits available to lend, the marginal resource cost of nondeposit-liability funds, and the marginal resource cost generated by allocating that dollar to lending, or $MC_L^i = r_N + C_N^i + C_L^i$. This schedule is the ibank's loan supply schedule. Summing across all banks' MC_L^i schedules yields the market supply of loan funds by the competitive banking industry, given by L_b^s in panel (b). The market loan rate, r_L^*, arises at the crossing point with the market demand schedule for loan funds across borrowers among the nonbank public (households and firms), denoted L_p^d. This loan rate, which corresponds to the bank's marginal revenue (MR_L^i) 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 amount at which marginal revenue equals marginal cost.

Banks are also suppliers of funds in the market for government securities. In panel (a) of Fig. 2.3, 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_S^i, equals the sum of the interest rate 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

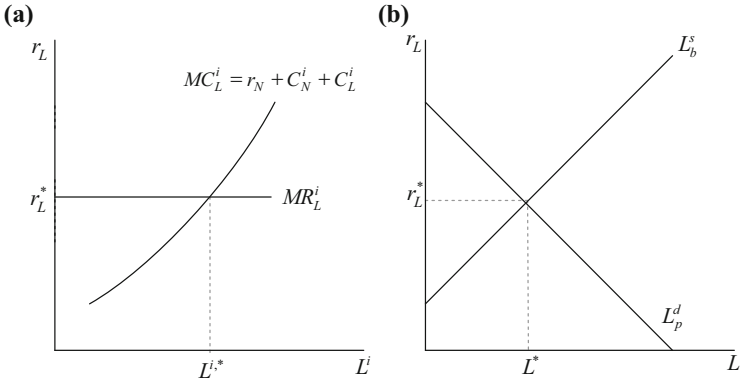

Fig. 2.2 A perfectly competitive bank loan market

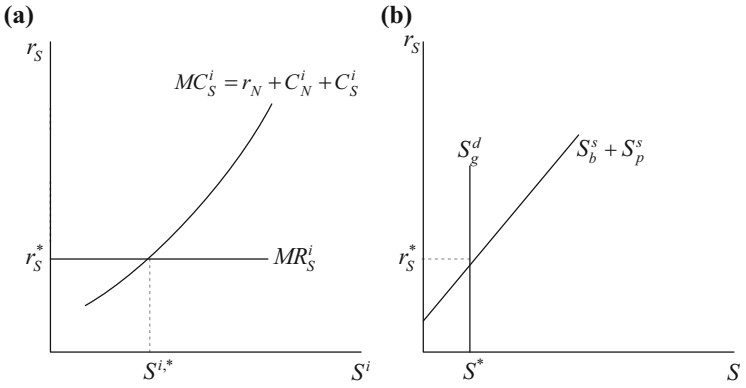

Fig. 2.3 The market for Government Securities

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. 2.3. 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 2.4 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, $(1-q)\left(r_S - C_S^i\right)$, and the marginal resource cost of deposits, C_D^i, or $NMR_D^i = (1-q)\left(r_S - C_S^i\right) - C_D^i$. 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. 2.4. At the crossing point with the nonbank public's market supply 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

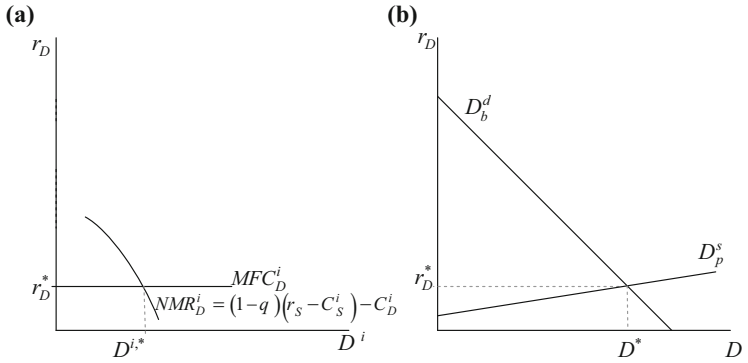

Fig. 2.4 A perfectly competitive bank deposit market

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, a bank takes the rate it must pay on funds raised by issuing nondeposit liabilities as given. As shown in panel (a) of Fig. 2.5, 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 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. 2.5. 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 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. 2.2, 2.3, 2.4 and 2.5 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 $r_S - C_S^i$, are equalized with total marginal interest and non-interest expenses on deposit and nondeposit liabilities, $(1-q)^{-1}\left(r_D + C_D^i\right)$ and $r_N + C_N^i$. Of course,

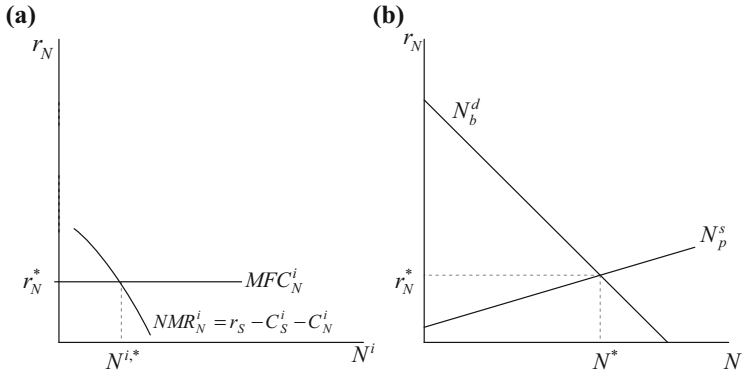

Fig. 2.5 The market for nondeposit-liability funds

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. 2.2. 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_L^i somewhat; nevertheless, $r_L - C_L^i$ remains higher than before. The bank therefore cuts back on security holdings, which reduces $r_S - C_S^i$ toward equality with the new, higher value of $r_L - C_L^i$. In addition, the bank raises more deposit and nondeposit-liability funds, which pushes up C_D^i and C_N^i and hence raises $(1-q)^{-1}\left(r_D + C_D^i\right)$ and $r_N + C_N^i$ to equality with this higher value of $r_L - C_L^i$. 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. 2.3, 2.4 and 2.5 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 $C_N^i \to 0$, 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:

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

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.

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 $C_N^i \to 0$ 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.

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 Elaysiani 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,

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

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:

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

According to this specific, separable implicit cost function, in which all parameters are nonnegative constants, banks face both contemporaneous costs governed by the $\alpha_1, \delta,$ and θ_1 parameters and intertemporal deposit and loan adjustment costs determined by the α_2 and θ_2 parameters.

A 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. Its contemporaneous loans depend on lagged, contemporaneous, and expected future interest rates on loans, deposits, and net interbank borrowings. The contemporaneous value-maximizing level of deposit funds demanded 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. 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

To consider imperfectly competitive behavior in banking markets, what are the economic implications, 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:

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

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

Figure 2.6 depicts determination of this bank's profit-maximizing loan and deposit rates and quantities of loans and deposits. 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 associated marginal revenue curve. It lends to the point at which marginal revenue, $\left[1 - (\varepsilon^i)^{-1}\right] r_L^i$, 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 $r_L^{i,M} = \left[1 - (\varepsilon^i)^{-1}\right]^{-1} = (r_N + C_N^i + C_L^i)$.

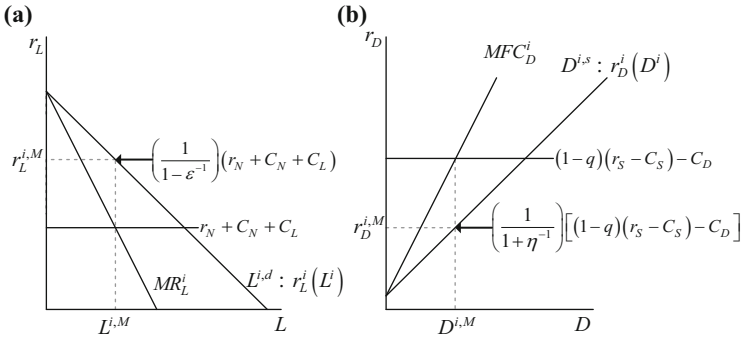

Fig. 2.6 Monopolistic loan and deposit markets

In panel (b) of Fig. 3.6, the bank is a monopsonist in the market for deposit funds. The bank issues deposits to the point at which the net marginal return from raising a dollar of deposits to allocate to securities holdings, $(1-q)\left(r_S - C_S^i\right) - C_D^i$, equals the marginal factor cost of rate of return on allocating the available portion of the marginal deposit dollar to securities, given by the quantity $\left[1 + (\eta^i)^{-1}\right] r_D^i$. 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 $r_D^{i,M} = \left[1 + (\eta^i)^{-1}\right]^{-1}\left[(1-q)\left(r_S - C_S^i\right) - C_D^i\right]$.

Of course, banks may lend or issue deposits in more than one market. Extending the model to include multiple loan and deposit markets would imply, as shown for instance by VanHoose (1985) and Hannon (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 $\frac{r_{L,1}^i}{r_{L,2}^i} = \frac{\left[1-\left(\varepsilon_2^i\right)^{-1}\right]\left(r_N + C_N^i + C_{L,1}^i\right)}{\left[1-\left(\varepsilon_1^i\right)^{-1}\right]\left(r_N + C_N^i + C_{L,2}^i\right)}$. The loan rate in market 1, therefore, is more likely to be higher than the loan rate charged in market 2 if the loan demand elasticity is lower and the marginal resource cost of lending is higher in 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 $\frac{r_{D,1}^i}{r_{D,2}^i} = \frac{\left[1+(\eta_2)^{-1}\right]\left[(1-q_1)\left(r_S - C_S^i\right) - C_{D,1}^i\right)\right]}{\left[1+(\eta_1)^{-1}\right]\left[(1-q_2)\left(r_S - C_S^i\right) - C_{D,2}^i\right)\right]}$. Thus, the deposit rate in market 1 is more likely to be lower than the deposit rate in 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 market 1 than in 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. 2.7, which is drawn under the simplifying assumption that $C_N^i = 0$, so that $r_S - C_S^i = r_N$, and hence the portfolio separation condition is 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, $C_{SS}^i, C_{LL}^i, C_{DD}^i \to 0$) 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. 2.7, 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.

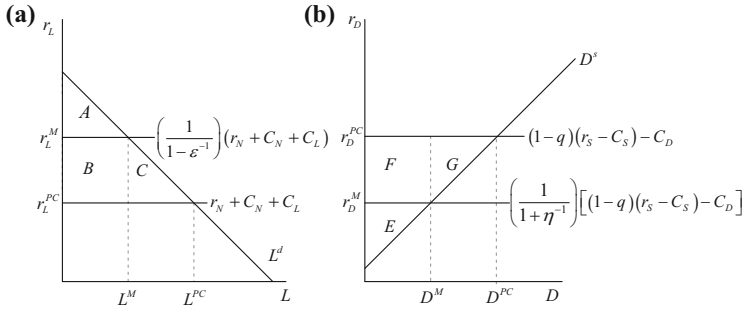

Fig. 2.7 Social losses with constant marginal resource costs

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. 2.7. 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.

Now 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 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. 2.7 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 with "large" numbers of rivals in bank loan and deposit markets, and hence low degrees of concentration. As shown by VanHoose (1988), this model can be extended to consider market entry to the point at which economic profits are extinguished.

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, $L_1^d(r_L^1, r_L^2)$, 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 $L_2^d(r_L^2, r_L^1)$, 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 loan demand given by $L^i = \frac{A_0}{m} + \frac{a}{m-1}\sum_{i \neq j}^{m}\left(r_L^j - r_L^i\right) - \frac{A_L}{m}r_L + \frac{A_S}{m}r_S$, 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 monposonistically competitive deposit market, given by $D^i = \frac{B_0}{n} - \frac{b}{n-1}\sum_{i \neq k}^{n}\left(r_D^k - r_D^i\right) + \frac{B_L}{n}r_D - \frac{B_S}{n}r_S$, 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 loan and deposit rates become more responsive to the security rate. In the limiting cases in which $m \rightarrow \infty$, the loan

rate becomes $r_L = r_S + C_L$, which in light of the simplifying assumptions is the perfectly competitive loan rate. For $n \to \infty$, 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), Cordella and Yeyati (2002), Dutkowsky and VanHoose (2013). 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 features of importance to potential customers. These attributes are distributed uniformly along a circle with a circumference equal to one, so 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 2.8 depicts a 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 $r_L^j + t\left(\frac{1}{n} - l_i^j\right)$.

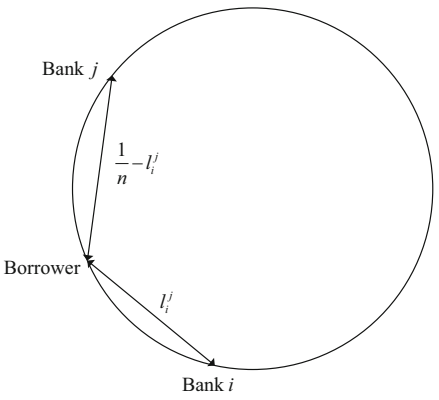

Fig. 2.8 Borrower and bank locations on a Salop circle

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 $l_i^j \leq \frac{1}{2n} + \frac{r_L^j - r_L^i}{2t}$. Of course, the same argument applies with respect to a borrower located on the opposite 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 $L^i = \left(l_i^j + l_i^k\right)L = \left[\frac{1}{n} + \frac{r_L^j + r_L^k - 2r_L^i}{2t}\right]L$. 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 $C_S \to 0$) profits of bank i are given by

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

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

$$\left[-t\frac{L^i}{L} + \frac{t}{n} + \frac{r_L^j + r_L^k}{2} - \frac{t}{L} - C_L\right] = r_S, \text{ and } -t\frac{D^i}{D} + \frac{t}{n} - \frac{r_D^j + r_L^k}{2} - \frac{t}{D} - C_D = (1-q)r_S.$$

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

$$r_L = r_S + C_L + \frac{t}{n}, \text{ and } r_D = r_S - C_D - \frac{t}{n}.$$

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, $\frac{t}{n}$. The amount of the markup is increasing in both the attribute cost and the market concentration of banks, measured by the reciprocal of the number of banks competing in the loan market. The monopsonistically competitive market deposit rate equals the perfectly competitive deposit rate, $r_S - C_D$, less a markdown, $\frac{t}{n}$. The magnitude of the markdown depends positively on the attribute cost and market concentration. Thus, given market concentration, 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.

References

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.

Alhadeff, David. 1954. *Monopoly and Competition in Commercial Banking*. Berkeley, CA: University of California Press.

Alon, Titan, John Fernald, Robert Inklaar, and J. Christina Wang. 2011. What is the value of bank output? Federal Reserve Bank of San Francisco Economic Letter 2011-15, May 16.

Basu, Susanto, Robert Inklaar, and J. Christina Wang. 2011. The value of risk: Measuring the service output of commercial banks. *Economic Inquiry* 49: 226–245.

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. 1965. Branch banking and economies of scale. *Journal of Finance* 20: 312–331.

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, and David Humphrey. 1997. Efficiency of financial institutions: International survey and directions for future research. *European Journal of Operational Research* 98: 175–212.

Besanko, David, and Anjan Thakor. 1992. Banking deregulation: Allocational consequences of relaxing entry barriers. *Journal of Banking and Finance* 16: 909–932.

Blair, Roger, and Arnold Heggestad. 1978. Bank portfolio regulation and the probability of bank failure. *Journal of Money, Credit, and Banking* 10: 88–93.

Baltensperger, Ernst. 1980. Alternative approaches to the theory of the banking firm. *Journal of Monetary Economics* 6: 1–37.

Chamberlin, Edward. 1962. *The Theory of Monopolistic Competition*. Cambridge, MA: Harvard University Press.

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.

Colwell, R.J., and E.P. Davis. 1992. Output and productivity in banking. *Scandinavian Journal of Economics* 94: S111–S129.

Cordella, Tito, and Eduardo Yeyati. 2002. Financial opening, deposit insurance, and risk in a model of banking competition. *European Economic Review* 46: 471–485.

Cosimano, Thomas. 1987. The federal funds market under bank deregulation. *Journal of Money, Credit, and Banking* 19: 326–339.

Cosimano, Thomas. 1988. The banking industry under uncertain monetary policy. *Journal of Banking and Finance* 12: 117–139.

Cosimano, Thomas, and John Van Huyck. 1989. Dynamic monetary control and interest rate stabilization. *Journal of Monetary Economics* 23: 53–63.

Dasgupta, Partha, and Joseph Stiglitz. 1981. Entry, innovation, exit: Towards a dynamic theory of industrial market structure. *European Economic Review* 15: 137–158.

Dewatripont, Mathias, and Jean Tirole. 1993. *The Prudential Regulation of Banks*. Cambridge, MA: MIT Press.

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.

Dutkowsky, Donald, and David VanHoose. 2013. Optimal bank sweeping and interest on demand deposits with imperfect competition and interest on reserves. *Journal of Macroeconomics* 38: 192–202.

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.

Flannery, Mark. 1982. Retail bank deposits as quasi-fixed factors of production. *American Economic Review* 72: 527–536.

Hancock, Diana. 1985. The financial firm: Production with monetary and nonmonetary goods. *Journal of Political Economy* 93: 859–880.

Hancock, Diana. 1991. *A Theory of Production for the Financial Firm.* Boston, MA: Kluwer.

Hannan, Timothy. 1991. Foundations of the structure-conduct-performance paradigm in banking. *Journal of Money, Credit, and Banking* 23: 68–84.

Hester, Donald, and James Pierce. 1975. *Bank Management and Portfolio Behavior*. New Haven, CT: Yale University Press.

Hotelling, Harold. 1929. Stability in competition. *Economic Journal* 39: 41–57.

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. Interest rate pass-through and financial crises: Do switching regimes matter? The case of Argentina. *Applied Financial Economics* 15: 77–94.

Kim, Daesik, and Anthony Santomero. 1988. Risk in banking and capital regulation. *Journal of Finance* 43: 1219–1233.

Klein, Michael. 1971. A theory of the banking firm. *Journal of Money, Credit, and Banking* 3: 205–218.

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.

Matutes, Carmen, and Xavier Vives. 2000. Imperfect competition, risk taking, and regulation in banking. *European Economic Review* 44: 1–34.

Miller, Stephen. 1975. A theory of the banking firm: Comment. *Journal of Monetary Economics* 1: 123–128.

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.

Perloff, Jeffrey, Larry Karp, and Amos Golan. 2007. *Estimating Market Power and Strategies*. Cambridge, UK: Cambridge University Press.

Pringle, John. 1973. A theory of the banking firm: A comment. *Journal of Money, Credit, and Banking* 5: 990–996.

Royster, Sara. 2012. Improved measures of commercial banking output and productivity. *Monthly Labor Review*, July, 3–17.

Salop, Steven. 1979. Monopolistic competition with outside goods. *Bell Journal of Economics* 10: 141–156.

Santomero, Anthony. 1984. Modeling the banking firm. *Journal of Money, Credit, and Banking* 16: 576–602.

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. 1985. Portfolio separation for stockholder-owned depository financial intermediaries. *Journal of Banking and Finance* 9: 477–490.
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.
Sheehan, Richard. 2013. Valuing core deposits. *Journal of Financial Services Research* 43: 197–220.
Startz, Richard. 1983. Competition and interest rate ceilings in commercial banking. *Quarterly Journal of Economics* 98: 255–265.
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.
VanHoose, David. 1985. Bank market structure and monetary control. *Journal of Money, Credit, and Banking* 17: 298–311.
VanHoose, David. 1988. Deregulation and oligopolistic rivalry in bank deposit markets. *Journal of Banking and Finance* 12: 379–388.

Chapter 3
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 ES theory suggests that cost 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 proposed bank mergers 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 (2004), Degryse and Ongena (2008), and Claessens (2009), 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 by considering the traditional SCP theory.

The Structure-Conduct-Performance Paradigm in Banking

Past expositions of the basic SCP paradigm as applied to the banking industry include Gilbert (1984), VanHoose (1984), and Hannan (1991). The essence of the SCP paradigm as applied to banking markets applies the basic oligopoly banking model discussed in the previous chapter.

© Springer-Verlag GmbH Germany 2017

D. VanHoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-662-54326-9_3

The SCP Hypothesis with Identical Banks

Chapter 2 showed that if there are m loan-market rivals and n competitors in deposit markets and if banks have identical, constant marginal resource costs associated with loans (L), deposits (D), and nondeposit liabilities (N), a Cournot-Nash outcome is the loan rate $r_L = \Omega(r_N - C_N - C_L)$, where $\Omega = \frac{m}{m-\varepsilon^{-1}}$ and ε is the loan demand elasticity. The 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 deposit supply elasticity.

The reasoning underlying the SCP hypothesis is depicted in Fig. 3.1, which revisits Fig. 2.7 in the prior chapter. In panel (a), if there are numerous loan-market competitors, so that $m \to \infty$ and $\Omega \to 1$, the loan rate is the perfectly competitive rate r_L^{PC}, and consumer surplus is the triangular region bounded beneath the 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, so that $m \to 1$ and $\Omega \to \frac{1}{1-\varepsilon^{-1}}$, the loan rate rises toward the monopoly level, r_L^M. Industry economic profits expand toward the maximum, monopoly level equal to the area of the rectangle denoted B, which is a transfer from consumers to banks. Consumer surplus shrinks toward the area of the triangle denoted A above the monopoly loan rate r_L^M. Lending falls toward the monopoly level L_M as the number of loan-market rivals shrinks, so 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 unattainable both to banks and borrowers.

In panel (b)—not drawn to the same scale as panel (a)—if there are numerous rivals in the deposit market, $n \to \infty$ and $\theta \to 1$, the deposit rate is at the perfectly competitive level r_D^{PC}, and depositors' producer surplus is the triangular region bounded above the 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, $n \to 1$ and $\theta \to \frac{1}{1+\eta^{-1}}$, and the deposit rate declines toward the monopsony level, r_D^M. The monopsonistic exploitation of depositors enlarges toward the maximum level equal to the area of rectangle F, which is a transfer from depositors to banks.

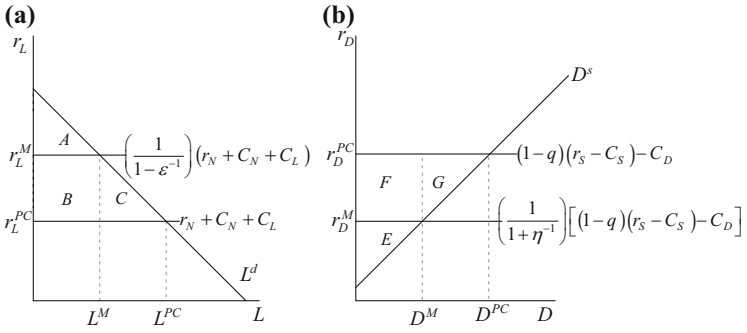

Fig. 3.1 The structure-conduct-performance hypothesis: Identical banks

Producer surplus shrinks toward the lower-left triangle denoted E. As the number of competitors in the deposit market shrinks, the deposits at banks decrease 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 generates higher loan rates and lower deposit rates. Loan and deposit quantities decline, which results in decreases in surpluses o and increases in deadweight losses.

Structural Asymmetry, Dominant Banks, and the SCP Paradigm

A symmetric market environment as 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 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. VanHoose (2013) provides an explicit model of dominant-bank behavior—and also see Balasubramanyan's (2014) policy application—but key elements of this approach can be visualized graphically.

A Dominant-Bank Model

Figure 3.2 provides an exposition of the dominant-bank framework as applied in a loan market. A large institution, bank i, and a number of smaller, "fringe," banks, indexed $j = 1,...m$, raise funds by issuing nondeposit as well as deposit liabilities. These fringe banks are assumed to face entry costs, so m is assumed to be relatively small. Both sets of institutions face upward-sloping marginal resource costs. Bank i's total marginal cost of lending expressed in terms of funds nondeposit funds is, as discussed in Chap. 2, given by $MC_L^i = r_N + C_N^i + C_L^i$, which is the upward-sloping curve graphed in panel (a) of Fig. 3.2. 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, it would face the loan demand curve L^d, its marginal revenue curve would be MR_M^i, its lending would be L_M^i and its loan rate would equal $r_L^{i,M}$. The bank's profits would equal 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.

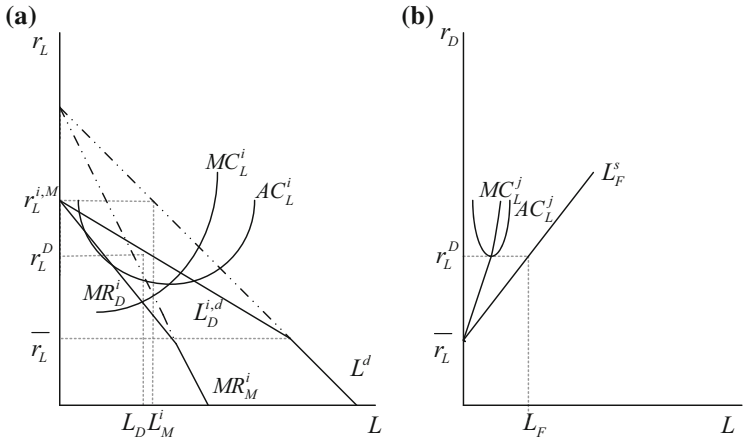

Fig. 3.2 The structure-conduct-performance hypothesis: A loan market with a dominant bank and a competitive fringe

Each of the m fringe competitors has less-efficient technologies for raising and lending funds. 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 $MC_L^j = r_N + C_N^j + C_L^j$ shown in panel (b) of Fig. 3.2. Each fringe banks stands willing to lend as long as it earns a loan rate exceeding a reservation loan rate, $\overline{r_L}$. 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 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 supply of loans of fringe banks is $L_F^s = \sum_{j=1}^{m} MC_L^j$ in panel (b).

In the face of competition from these fringe institutions, above their reservation loan rate $\overline{r_L}$, bank i confronts not the entire market demand curve L^d in panel (a) but instead the *residual* demand curve $L_D^{i,d} = L^d - L_F^s$. At each possible loan rate above $\overline{r_L}$, loans demanded from the "dominant" institution, bank i, equal the market quantity demanded less the amount supplied by all fringe banks at that rate. Associated with this residual demand curve is the marginal revenue curve MR_D^i. Bank i maximizes its profits by reducing loans to L_D^i. The profit-maximizing loan rate is r_L^D, and bank i's maximum profit 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, their lending equals 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 panel (b), each fringe bank j earns zero economic profits, implying that there is no incentive for fringe banks either to enter or exit the market.

Thus, Fig. 3.2 shows a long-run equilibrium, prior to which there might have been fewer fringe banks and hence a less elastic aggregate fringe supply curve, a higher profit-maximizing loan rate for the dominant bank, and positive profits for fringe banks that would have encouraged entry by additional banks. Fringe banks earn no economic profits, but the dominant bank earns positive (but non-maximized) profits due to its cost advantage.

Strategic Entry Deterrence

If fringe entry costs are low, then at some point following entry, one or more fringe rivals may, in the process of learning by doing, discover how to replicate the technology that provides bank i with an edge. In a dynamic setting, recognition of this fact 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. 3.2. As long as this loan rate exceeds average cost at its profit-maximizing lending, bank i could thereby forestall entry by fringe banks and ensure 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 a sufficient concern that a fringe rival could replicate its technology. If so, the expected discounted stream of profits under this entry-deterrence strategy exceeds a strategy that permits entry.

Evaluating the Applicability of the SCP Paradigm to the Banking Industry

The SCP hypothesis suggests that in more concentrated banking markets, *ceteris paribus*, 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 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? Traditionally, the most common measures of bank market concentration used to address this question 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-Hirschmann 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, the preponderance of which was generally supportive. Higher bank market concentration appeared to generate higher loan rates, lower deposit rates, and increased industry profits. Indeed, Rhoades (1982) concluded from 1969–1978 U.S. banking data that monopoly power in banking reduced loans to individuals by about 16% and pushed up profits by approximately 13%. In a more recent evaluations of SCP predictions for loan rates, Shaffer and Srinivasan (2002) have found strong evidence of higher loan rates resulting from increased concentration in data samples ranging from 2500 to 3900 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 a study of data from more than 200 Spanish banks in about 50 geographic markets, Martín-Oliver et al. (2008) find 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. In a study of more than 7000 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 Hannon 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. 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. More recently, Montgomery et al. (2014) examine Japanese bank data during the 1996–2006 interval and find significant cost inefficiencies associated with market concentration but, consistent with the SCP hypothesis, higher profit efficiency made possible by greater market power. Brewer and Jackson (2006) account for risk in the derived demand for deposits and find evidence that the effect of market concentration on deposit rates is about 50% smaller than

corresponding estimates that fail to account for risk. They conclude that traditional cross-sectional SCP studies likely overstate the influence of concentration on loan and deposit rates.

Evidence from Cross-Country Studies

Although much evidence in favor of the SCP hypothesis has focused on U.S. banking, a few 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. Marrouch and Turk-Ariss (2014) examine data about 100 developing nations in which core deposits are key funding sources for loans and in which lending and deposit-taking functions are concentrated among relatively small numbers of banks. Consistent with the SCP theory, they find evidence of significant economic rents accruing to these banks. In addition, in an examination of the determinants of fees charged by banks in 703 banks in 46 countries in 2011, Tennant and Sutherland (2014) find that market concentration contributes to bank profits from charging fees.

Dynamic Interest Rate Responses: Competition and Pass-Through Effects

Since the early 1990s, studies have broadened the scope of inquiry beyond cross-sectional analysis to consideration of time-series evidence. 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 imperfect5 competition. Hannon and Liang (1993) have also considered the time-series relationship between securities rates and bank rates in an effort to infer a relationship between retail market concentration and 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

deposit-rate determination, and Winker (1999), who suggests that asymmetric information and credit risks may also play some role in incomplete pass through. In an examination of deposit-rate determination in U.S. banking markets, Rosen (2007) concludes that there is evidence of a complex dynamic interplay between the 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 (2012). 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—an essential prediction for which Banerjee et al. (2013) separately find empirical support. In Kopecky and VanHoose's theoretical model, greater market concentration enables banks to restrain loan and deposit adjustments—a prediction for which Adams and Amel (2011), Fungàčovà et al. (2014), and Leroy (2014) separately provide empirical support—and thereby impose larger loan rate markups on borrowers and greater deposit rate markdowns on depositors—there is less complete 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 contemporaneously more fully alongside the competitive securities rate, as predicted by standard static models of perfectly competitive banking markets. Also see Abo-Zaid (2015), Dia (2013), and Dia and Giuliodori (2012) for similarly structured dynamic banking frameworks with imperfect competition.

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 Argentina (Humala-Acuna 2005), Austria (Burgstaller 2005), Canada (Scholnick 1999), Chile (Bernstein and Fuentes 2004), the Eurozone (Leroy and Lucotte 2015; Hristov et al. 2014; Hülsewig and Mayer 2009; Vajanne 2009), Finland (Kauko 2005), Germany (Winker 1999), Malaysia and Singapore (Scholnick 1996), New Zealand (Liu et al. 2011), Turkey (Aydin 2007), Norway (Raknerud et al. 2011), the United Kingdom (Heffernan 1997; Fuertes and Heffernan 2009; Fuertes et al. 2010; and Harimohan et al. 2016), and the United States (Hannon and Liang 1993; Jackson 1997; Kahn et al. 2005; Neumark and Sharpe 1992; Sheehan 2006) and studies examining broader cross-country evidence, including Aspergis and Cooray (2015), 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), Sørenssen and Werner (2006), Tieman (2004), Toolsema et al. (2001), van Leuvensteijn et al. (2013), and Amidu and Wolfe (2013). These studies have reached mixed conclusions. Some studies, such as Neumark and Sharpe (1992), Scholnick (1999), and Craig and Dinger (2010) 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. Belke et al. (2013), who utilize nonlinear cointegration techniques in analysis of 2003–2011 interest-rate data for a dozen euro nations, find evidence of less pass through for household loans than for business loans.

One finding common to nearly every study is evidence of sluggish and incomplete pass through from market rates to bank retail rates. In virtually every country examined, changes in security 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 in Cottarelli and Kourelis (1994), who estimated contemporaneous impacts of market rates on bank rates from as low as 0.06 to not much above 0.80. Thus, time-series pass-through studies are generally supportive of SCP 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 predicts that a situation with asymmetric competition depicted Fig. 3.2 cannot exist unless a dominant bank possesses a technological edge over its fringe 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. Also, if entry costs faced by fringe banks are substantial, existence of a cost advantage enables a dominant banks to set its loan rate independently and experience positive economic profits even when some fringe entry occurs.

Consistent with the basic implications of the dominant-bank model, Craig and Hardee (2007) find 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 dominant-bank 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 a study of rural-county banking markets between 1996 and 2007, Cyree and Spurlin (2012) also find that large banks in these markets earn non-competitive rents. In addition, they conclude that small banks can compete because of profit inefficiencies that exist at those large banks.

In deposit markets, Hannan (2006), Hannan and Prager (2004, 1998), Park and Pennacchi (2008) find pricing behavior by large, multimarket banks and smaller, single-market banks 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 afterward have dissipated. Hannan and Prager (2009) conclude that profitability of small banks in rural markets depends on whether these banks operate in single markets and whether competing banks are multi-market institutions. Finally, DeYoung (2003) finds 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.

The dominant-bank theory predicts that fringe banks 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. In light of both the smaller scale of lending by a fringe bank and its higher average cost, fringe banks' profitability should be lower than the profitability of dominant banks. Consistent with these predictions, Amel and Liang (1997) find evidence in U.S. data from the late 1970s through the late 1980s that bank entry responds to positive profits. They conclude that profits of incumbent large and particularly small banks decline in response to such entry.

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 the dominant bank's short-run profits but 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. Adams and Amel's (2016) finding, based on a study of 3000 urban and rural U.S. banking markets between 1994 and 2008, that greater market concentration is associated with entry could be viewed as consistent with such the use of such a strategy by dominant banks. So is their finding of a direct relationship between past and current entry. Thus, arguably consistent with effective entry deterrence by dominant banks, a lower extent of prior entry tends to be associated with less current entry. They note that most new entry into local U.S. banking markets in recent years has occurred via branching by existing banks; indeed, as Adams and Gramlich (2016) document, virtually no de novo entry has occurred in U.S. markets since the recent banking crisis and the subsequent low-interest-rate environment.

Not all empirical evidence supports 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 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 5000 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. Barbosa et al. (2015) utilize the Panzar-Rosse approach in a study of Brazilian banking data and argue that it may be biased against indicating stronger market power when banks offer wider arrays of products. Goddard and Wilson (2009) also suggest that this approach requires a dynamic formation of the revenue relationship to avoid a potential misspecification bias. In addition, Bikker et al. (2012) caution that an unscaled revenue function should be used to infer information about the degree of competition using the Panzar-Rosse methodology.

Market Structure and Bank-Customer Relationships

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. 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." As discussed in detail by Elyasiani and Goldberg (2004) and highlighted

below, however, theoretical implications and empirical evidence regarding customer relationships and their effects are somewhat mixed.

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 $L_t^{i,d}\left(r_{L,t}^i; L_{t-1}^i, L_{t-2}^i, \ldots L_{t-m}^i; X_t^i\right)$, and confronts the deposit supply function $D_t^{i,s}\left(r_{D,t}^i; D_{t-1}^i, D_{t-2}^i, \ldots D_{t-m}^i; Y_t^i\right)$, 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

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

where R is the bank's subjective discount factor, with $0 < R < 1$, and, to further simplify, resource costs are assumed to be negligible. This assumption implies portfolio separation, as discussed in Chap. 2—an assumption that receives some support, at least in terms of spillover effects resulting from imperfect competition, from the findings obtained by Adams et al. (2002). 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

$$\left[1-\left(\varepsilon^{i}\right)^{-1}\right]r_{L,1}^{i}=r_{S,1}-R\frac{\partial L_{2}^{i,d}}{\partial L_{1}^{i}}E_{1}\left(r_{L,2}^{i}-r_{S,2}\right),\ \text{and}$$

$$r_{D,1}^{i}\left[1+\left(\eta^{i}\right)^{-1}\right]=(1-q)r_{S,1}+R\frac{\partial D_{2}^{i,s}}{\partial D_{1}^{i}}E_{1}[(1-q)r_{S,2}-r_{D,2}],$$

where, as in Chap. 2, ε^{i} is the loan demand elasticity and η^{i} is the deposit supply elasticity.

If $\frac{\partial L_{2}^{i,d}}{\partial L_{1}^{i}}=0$ and $\frac{\partial D_{2}^{i,d}}{\partial D_{1}^{i}}=0$, so that there are no holdover effects from bank-customer relationships going into the second period, the above conditions choices reduce to conditions like those discussed in Chap. 2. If, however, $0<\frac{\partial L_{2}^{i,d}}{\partial L_{1}^{i}}<1$ and $0<\frac{\partial D_{2}^{i,d}}{\partial D_{1}^{i}}<1$, customer relationships matter, and the terms $R\frac{\partial L_{2}^{i,d}}{\partial L_{1}^{i}}E\left(r_{L,2}^{i}-r_{S,2}\right)$ and $R\frac{\partial D_{2}^{i,s}}{\partial D_{1}^{i}}E_{1}[(1-q)r_{S,2}-r_{D,2}]$ 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. van der Cruijsen and Diepstraten (2015) discuss, in the context of data from the Netherlands, how customer predispositions not to switch away from existing banking relationships can give rise to such intertemporal interactions.

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 customer relationships. 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 current profits for greater future profits. 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. Current profits of incumbents will be lower, reducing the incentive for entry to occur and resulting in a more concentration, *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—an outcome consistent with experimental evidence provided by Cornée et al. (2012)—which ultimately could lead to more entry. Wood's model indicates, though, 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.

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. Cotugno et al. (2013), in a study of more than 5000 Italian bank-lending relationships with firms, find that banks also are more likely to extend credit to borrowers with which they have longstanding relationships. An analysis of loans by a Bulgarian bank between 2003 and 2007 by Kirschenmann (2016) finds greater credit rationing of more opaque firms but that the degree of rationing diminishes with extended lending relationships with firms.

Hannan and Adams (2011) explore 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 and Adams find 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 and Adams find, 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.

More empirical evidence recently has emerged regarding the nature of information that banks glean from customer relationships. Botsch and Vanasco (2015) focus on trying to infer the presence of signals to banks from customer relationships not available to other participants in financial markets, and based on 1987–2003 data from 7618 syndicated loans conclude that a considerable portion of bank information acquisition from these relationships is private. Chang et al. (2014) examine a proprietary-data sample of more than 25,000 loans involving in excess of 3600 Chinese firms and find evidence that information that banks derive from relationships improves default predictions, particularly for lengthier relationships. Separate results obtained from over 2600 finance-company loans to nearly 1500 Taiwanese firms by Chen et al. (2015) suggest that "soft" information commonly available in relationship lending helps in predicting loan defaults. In addition, analysis of data involving bank loans to 43,000 Italian firms between 2008 and

2010 leads Fiordelisi et al. (2014) to conclude that the likelihood of loans becoming distressed declines with strengthened lending relationships. Fredriksson and Moro (2014) find evidence, derived from a study of more than 4000 Finnish business loans during the 2001–2005 interval, that customer relationships contribute to the bank-profitability-enhancing effects of better borrower performance. Furthermore, Li et al. (2015) examine a large set of U.S. syndicated loans, which they categorize as either "relationship" or "transactional" lending, and conclude that the former experience less ex post quality deterioration, which they credit to stronger monitoring of relationship lending. Finally, based on analysis of more than a million German loans, Puri et al. (2011) conclude that customer relationships provide screening and monitoring benefits that significant reduce the likelihood of loan defaults.

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 a key feature of relationship banking is access to inelastically supplied core 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. Analysis of loan-rate spreads for more than 80,000 Italian firms between 2008 and 2010 conducted by Gambacorta and Mistrulli (2014) suggests that borrowers with stronger bank lending relationships also pay lower spreads.

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 by 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 3000 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 impose tougher loan terms on borrowers with whom they have maintained a continuing relationship. Xu et al. (2015) provide evidence, based on more than 9000 loans extended by more than 600 Chinese banks, that banks in imperfectly competitive markets can apply informational gains from customer relationships to extract rents via collateral requirements—which it is important to note that Hainz et al. (2013) find to be inversely related to the extent of bank-market

competition. In addition, Calomiris and Pornrojnangkool (2009) find in a study of more than 7000 jointly produced loan and security underwritings that depository institutions charge premiums for both lending and underwriting services, although the terms at which they provide these services are better than those of investment banks.

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 of 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—which Petersen and Rajan (2002) and DeYoung et al. (2011) document already have shrunk over time with respect to small-firm borrowers with greater bank utilization of information technologies aimed at "hardening" information—on loan rates disappear. This fact, Agarwal and Hauswald suggest, indicates the main effect of distance is its impact on banks' capabilities to acquire and take advantage of "soft" information in relationships with borrowers—information that Hattori et al. (2015) conclude, on the basis of a survey of Japanese bankers, primarily is collected by branch managers. 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. Berger and Black (2011) offer evidence that larger firms with loans involving "soft" information are indeed the main source of a relationship advantage for smaller banks with fewer management layers. Based on a study of a stratified-sample survey of Italian firms, Cenni et al. (2015) conclude that larger firms with longer relationships also are less likely to be denied credit, particularly if most of their borrowing is concentrated with the bank with which they have a relationship.

Competition and Relationship Lending

How does increased competition in loan markets affect relationship lending? On the one hand, based on a model emphasizing firms' internalization of benefits of lending relationships, Petersen and Rajan (1995) argue that the such benefits 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 3000 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 (2007, 2008) 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 4000 Italian firms, Presbitero and Zazzaro (2011) suggest that these results may be driven partly by size and distance effects rather than by the level of bank market concentration, per se. They also argue that the direction of competitive effects on relationship lending hinges on whether large, dominant banks are present or absent from a banking market.

Ergungor's (2005) examination of the effects of small business lending on U.S. bank performance between 1996 and 2002 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. Additionally, De la Torre et al. (2010) conclude that the recent intensification of competition in lending to smaller businesses in many developing nations has occurred in the absence of relationships. Finally, Beck et al. (2011) find, based on analysis of data for 91 banks based in 45 nations, that characteristic of loans to small and medium-sized enterprises are not strongly correlated either with lending technologies or with bank organization structures, a result that they argue reduces the likely importance of relationship lending. 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 ts 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, at times larger institutions could operate with higher profits. 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 4000 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 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 several 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% in the United States) or not publicly traded (nearly 90% in the United States). Although Altunbas et al. (2001), however, find no evidence of systematic differences in cost efficiency across mutual, private, and publicly traded banking institutions, Delis and Tsionas (2009) find evidence of a negative relationship between market power and efficiency. 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% lower than at the best-practice banks, or slightly above the 20–25% 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 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. Furthermore, based on an analysis of global banks with more than $50 billion in assets, Davies and Tracey (2014) contend that once

implicit subsidies associated with regulatory "too-big-to-fail" guarantees (see Chap. 6) are taken into account, evidence of scale economies vanishes.

Hughes and Mester (1998, 2013) 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 likewise conclude that banking is characterized by significant scale economies once financial capital structure and risk are taken into account. Hughes and Mester (2013) likewise find evidence of scale economies after controlling for risk in a study of data on the largest U.S. bank holding companies in 2003, 2007, and 2010. Separately, Beccalli et al. (2015) examine a dozen years of data from banks encompassed by the European Stoxx 600 Banks index and find evidence favoring scale economies for all bank size categories examined.

Wheelock and Wilson (2012) examine data on all U.S. commercial banks between 1984 and 2006 and a subset of commercial banks between 2001 and 2006 and find evidence of increasing returns to scale. They conclude that irrespective of other contributors to U.S. banking consolidation, economies of scale is an unambiguous factor.

Evidence on scope economies is 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.

Examination of 2004–2006 proprietary data on a large Italian bank's loan portfolio leads Bellucci et al. (2013) to conclude that loan terms vary as marginal cost changes with distance-related informational factors. In contrast to Degryse and Ongena (2005), therefore, Belucci et al. conclude that, consistent with the efficient-structure hypothesis, varying loan terms reflect price differentiation rather than spatial price discrimination that the SCP theory would predict. Homma et al. (2014) utilize 1974–2005 data for Japanese banks to conclude that, consistent with the efficient-structure approach, the most cost-efficient banks generally emerge as predominant, although they also find evidence that increased market concentration tends to reduce efficiency levels.

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 (1989, 1993) 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.

Divergences across countries, in contrast, is the theme of work by Gonzalez (2009), who analyzes 1996–2002 banking and political economy data from 69 nations. Gonzalez generally finds support for the efficient-structure theory. Nevertheless, country-specific differences in bank entry requirements, deposit insurance, and other institutional elements influence the degree of support. Hsieh and Lee's (2010) study of banking data from 61 countries between 1992 and 2006 focuses on the relationship between bank competition and profits and likewise finds that political economy variables are key elements conditioning the nature of that relationship. Mirzaei and Moore (2014) conclude from analysis of data from 146 nations between 1999 and 2011 that the intensity of bank competition also hinges on political economy elements.

Dai and Yuan (2013) evaluate the social-efficiency effects of bank entry depending on the degree of product differentiation for new entrants vis-à-vis incumbents. Based on their analysis of U.S. banking markets between 2000 and 2008, they conclude that bank retail markets are highly segmented and that little evidence exists of situations of over- or under-entry. Dai and Yuan also argue that their results indicate that most efficiency gains from bank entry occur as a consequence of a broadening of product variety. Indeed, one finding is that entry of banks with the same attributes as incumbent banks is associated with a decline in consumer welfare.

In some cases, authors conclude that behavior in certain banking markets entails Bertrand-Nash price-rivalry behavior among oligopolistic rivals offering heterogeneous products. Barros (1999) applies a Salop-style model (see Chap. 2) 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 considered 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.

Much of the research on the efficient structure approach focuses on implications of scale economies for retail loan and deposit rates and quantities. Bos et al. (2013) examine how market structure influences innovative activity of banks. Using 1984–2004 U.S. data, they find evidence of an inverted-U-shaped relationship between competition and innovation and conclude that recent structural changes in the U.S. banking industry have generated too little innovative activity in banking. More research on this "dynamic" issue appears warranted, whether from an SCP or an efficient-structure perspective, particularly given evidence provided by Duygun et al. (2013) indicating that long-run bank cost and profit efficiency is related to banks' innovative activities, by He (2015) and Hernández-Murillo et al. (2010) that banks engage in strategic innovations competitively, and by Feng and Serletis (2010) suggesting that total factor productivity in the banking industry has been declining.

The efficient-structure theory also presupposes the existence of institutions that promote competitive behavior and thereby allow market entry in banking to enhance the intensity of competition even at relatively large scales. Based on an analysis of the interrelationships between the degree of bank market power and measures of institutional development, Delis (2012) concludes that sufficient institutional development is prerequisite for attaining competitive and efficient banking markets. In a cross-country study of profit persistence at more than 11,000 banks in 65 nations, Goddard et al. (2011) reach an analogous conclusion.

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 rise because of outlays to boost product demands, however, their profits fall, which removes the incentive for additional rivals to enter. In Sutton's framework with vertical product differentiation and endogenous sunk fixed costs, therefore, steady increases in market size ultimately generate no further increases in the number of firms, and concentration eventually reaches a lower bound.

Endogenous Sunk Costs and Concentration

To review the essential elements of Sutton's theory, consider this adaptation of the stylized exposition by Belleflamme and Peitz (2015) to the banking industry; see also Shiman (2007) for an arguably more general exposition. Consumers spend a fraction γ of their income, denoted y, on the flow of services offered by a particular financial industry—presumed in this case to be the banking industry that is assumed to engage in Cournot rivalry in its loan market. The typical consumer's a service-flow quantity is denoted l, and a consumer receives utility $u(l_0, l) = l_0^{1-\gamma}(ul)^{\gamma}$, where u is the quality of the banking industry's service flow—assumed here to be the services provided by lending—and l_0 is the quantity of an outside industry's financial service flow. Suppose that the mass of banking consumers is M. Then total consumer expenditure—and hence total industry revenue—is $R = M\gamma y$, and for any two banks i and j among the n incumbent banks in the industry, it must be true that the ratio of loan rates to service-flow quality is equalized, so that $\frac{r_L^i}{u_i} = \frac{r_L^j}{u_j} = \lambda$, which in turn implies that industry revenues are equal to $R = \sum_{i=1}^{n} r_L^i l_i = \lambda \sum_{i=1}^{n} u_i l_i$.

Belleflamme and Peitz show that in a two-stage-game with an entry cost e and an exogenous sunk fixed cost of delivering a quality level u_i given by $C(u_i)$, in a symmetric Cournot equilibrium with $u_i = u$ for all i, the net profit of one of these stylized banking firms facing constant per-unit variable cost c is equal to $\frac{M\gamma y}{n^2} - e - C(u)$. Hence, as the mass of consumers increases in magnitude—that is,

as the size of the banking market expands—so will the number of banking firms n in a zero-profit equilibrium for an industry with exogenous sunk fixed costs.

Belleflamme and Peitz show that if a given market incumbent i's endogenous sunk fixed costs are determined according to $C(u_i) = \alpha u_i^{\beta}$, in advance of entry the variable profits that incumbent i would anticipate earning post-entry are given by $\left(1 - \frac{n-1}{u_i \sum_{j=1}^{n} \frac{1}{u_j}}\right)^2 R - \alpha u_i^{\beta}$, which yields as the optimal quality choice in a symmetric equilibrium for the n firms that become industry incumbents:

$$\bar{u} = \left(\frac{M\gamma y(n-1)^2}{\alpha\beta n^3}\right)^{1/\beta} \tag{3.1}$$

Thus, the Cournot-equilibrium quality depends positively on market size and the share of consumer income spent on the bank product. This quality level depends negatively on the number of banks in the industry and the parameters governing the magnitude of endogenous sunk fixed costs of generating quality. The equilibrium quality choice yields a post-entry profit for an incumbent equal to $\frac{M\gamma y}{n^3}\left[n - \left(\frac{2}{\beta}\right)(n-1)^2\right] - e$, which can feasibly be positive only if $n - \left(\frac{2}{\beta}\right)(n-1)^2 > 0$, a condition that is independent of the market size. Consequently, there is an upper bound on the number of firms given by $n \leq 1 + \frac{1}{4}\left\{\beta + [\beta(\beta+8)]^{\frac{1}{2}}\right\} = \bar{n}$, and hence a lower bound on industry concentration.

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

Sutton's framework applies to industries in which non-price competition is an essential feature. It has long been recognized 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 VanHoose (1988) discusses, however, implicit deposit rates measured as remitted service charges arise in imperfectly competitive banking markets with or without deposit rate ceilings and dissipate in perfectly competitive markets in which service charges adjust to reflect marginal costs of providing relevant services. As noted by Michael Klein (1978), such implicit rates are measured ex post and cannot function as market signals for depositors, who lack knowledge of bank costs to compute implicit rates of interest. Indeed, customers do not care about bank costs per se but base their decisions on the characteristics of banks' products.

Thus, a more productive approach to contemplating non-price 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 that quality of service—such as weekly office hours of walk-in or drive-in services and availability of 24-h automated services derived from survey data—influences behavior of bank customers and found an inverse relationship between bank concentration and service quality levels. White (1976) found a similar result using the number of branches as the relevant quality variable in U.S. banking data from 1970. Carlson and Mitchener (2006) suggest that 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 1989–1990 and find that although branch banking helps differentiate banks' products, 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 1988–1995 Norway data, 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 are few studies on the economic impacts of advertising in the banking industry. Martín-Oliver and Sals-Fumás (2008), who examine 1983–2003 data for Spanish banks, conclude that advertising outlays succeeded in boosting deposit supply and loan demand, 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 market concentration and advertising outlays and each contradictory conclusions. 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, as do the more recent and broader studies by DeYoung and Örs (2004) and Örs (2006). DeYoung and Örs examine data from almost 1900 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 1994–2000 data from nearly 4500 commercial banks 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 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 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. Hasan and Smith (1997, p. 48) speculate that "it is possible that existing banks use. . .fixed costs to deter entry and thereby raise profits." Furthermore, Dell'Ariccia (2001) builds on Dell'Ariccia et al. (1999) to provide a theory of endogenous sunk fixed costs generated by adverse selection. 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.

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 fewer than 100,000 people to more than two million. She finds little variation in bank market concentration across different markets sizes, as shown in Fig. 3.3, 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 1850. Dick shows that scatterplots relating either 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.

Cohen and Mazzeo (2010) conduct an analysis complementary to Dick's that focuses on non-metropolitan-statistical-area "labor market areas" as defined by the U.S. Bureau of Labor Statistics. These researchers emphasize bank branches as fixed investments in the form of sunk costs and utilize branching density as a key explanatory variable. They offer evidence that banks engaging in multimarket competition via branching induce branching investment responses when they penetrate markets with new branches. The results, they argue, are endogenous adjustments in market structure that typically result in increased market concentration.

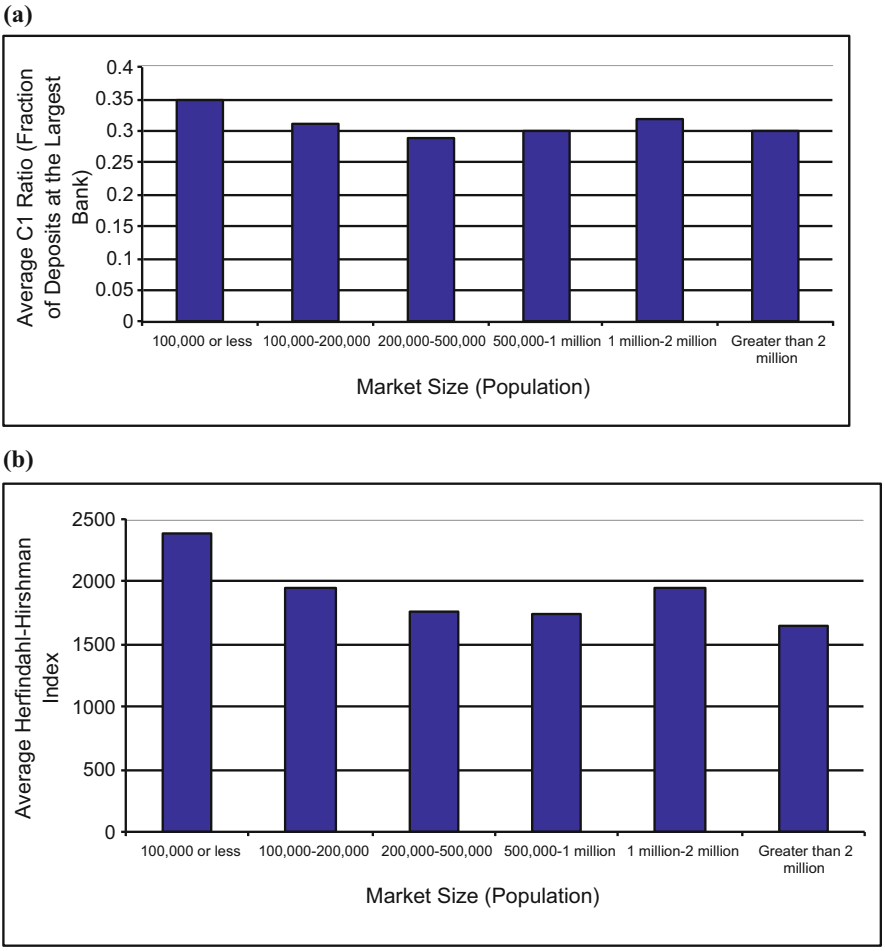

Fig. 3.3 The relationship between U.S. regional bank market concentration and market size [Source: Dick (2007)]

References

Abo-Zaid, Salem. 2015. Optimal monetary policy with the cost channel and monopolistically competitive banks. *Journal of Macroeconomics* 45(C): 284–299.

Adams, Robert, and Dean Amel. 2011. Market structure and the pass-through of the federal funds rate. *Journal of Banking and Finance* 35: 1087–1096.

Adams, Robert, and Dean Amel. 2016. The effects of past entry, market consolidation, and expansion by incumbents on the probability of entry in banking. *Review of Industrial Organization* 48: 95–118.

Adams, Robert, and Jacob Gramlich. 2016. Where are all the new banks? The role of regulatory burden in new charter creation. *Review of Industrial Organization* 48: 181–208.

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.

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.

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.

Allen, Linda, and Anoop Rai. 1996. Operational efficiency in banking: An international comparison. *Journal of Banking and Finance* 20: 655–672.

Allen, Linda, Anthony Saunders, and Gregory Udell. 1991. The pricing of retail deposits: Concentration and information. *Journal of Financial Intermediation* 1: 335–361.

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.

Amidu, Mohammed, and Simon Wolfe. 2013. The effect of banking market structure on the lending channel: Evidence from emerging markets. *Review of Financial Economics* 22: 146–157.

Anand, Bharat, and Alexander Galetovic. 2006. Relationships, competition, and the structure of investment banking markets. *Journal of Industrial Economics* 54: 151–199.

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.

Aspergis, Nicholas, and Arusha Cooray. 2015. Asymmetric interest rate pass-through in the U.S., the U.K., and Australia: New evidence from selected individual banks. *Journal of Macroeconomics* 45: 155–172.

Aydin, Halil. 2007. Interest rate pass-through in Turkey, Central Bank of the Republic of Turkey.

Bain, Joseph. 1956. *Barriers to New Competition*. Cambridge, MA: Harvard University Press.

Balasubramanyan, Lakshmi. 2014. Differential capital requirements: Leverage ratio versus risk-based capital ratio from a monitoring perspective. Federal Reserve Bank of Cleveland Working Paper 14–15, October.

Banerjee, Anindya, Victor Bystrov, and Paul Mizen. 2013. How do anticipated changes to short-term market rates influence banks' retail interest rates? *Journal of Money, Credit, and Banking* 45: 1375–1414.

Barbosa, Klenio, Bruno de Paula Rocha, and Fernando Salazar. 2015. Assessing competition in the banking industry: A multi-product approach. *Journal of Banking and Finance* 50: 340–362.

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.

Beccalli, Elena, Mario Anolli, and Giuliana Borello. 2015. Are European banks too big? Evidence on economies of scale. *Journal of Banking and Finance* 58: 232–246.

Beck, Thorsten, Aslie Demirgüç-Kunt, and Maria Soledad Martinez Pería. 2011. Bank financing for SMEs: Evidence across countries and bank ownership types. *Journal of Financial Services Research* 39: 35–54.

Becker, William. 1975. Determinants of the United States currency-demand deposit ratio. *Journal of Finance* 27: 57–74.

Belke, Ansgar, Joscha Beckmann, and Florian Verheyen. 2013. Interest-rate pass through in the EMU: New evidence from nonlinear cointegration techniques for fully harmonized data. *Journal of International Money and Finance* 37: 1–24.

Belleflamme, Paul, and Martin Peitz. 2015. *Industrial Organization: Markets and Strategies*. Cambridge, UK: Cambridge University Press.

Bellucci, Andrea, Alexander Borisov, and Alberto Zazzaro. 2013. Do banks price discriminate spatially? Evidence from small business lending in local credit markets. *Journal of Banking and Finance* 37: 4183–4197.

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.

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.

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.

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. 2003. The economic effects of technological progress: Evidence from the banking industry. *Journal of Money, Credit, and Banking* 35: 141–176.

Berger, Allen, and Lamont Black. 2011. Bank size, lending technologies, and small business finance. *Journal of Banking and Finance* 35: 724–735.

Berger, Allen, and Timothy Hannan. 1989. The price-concentration relationship in banking. *Review of Economics and Statistics* 71: 291–299.

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 David Humphrey. 1991. The dominance of inefficiencies over scale and product mix economies in banking. *Journal of Monetary Economics* 28: 117–148.

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, 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 Gregory Udell. 1995. Relationship lending and lines of credit in small firm finance. *Journal of Business* 68: 351–381.

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, 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, 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, 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.

Berlin, Mitchell, and Loretta Mester. 1999. Deposits and relationship lending. *Review of Financial Studies* 12: 579–607.

Berstein, Solange, and J. Rodrigo Fuentes. 2004. Is there lending rate stickiness in the Chilean banking industry? In *Banking Market Structure and Monetary Policy*, ed. Luis Antonio Ahumada and J. Rodrigo Fuentes, 183–210. Santiago: Central Bank of Chile.

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.

Bikker, Jacob, Sherrill Shaffer, and Laura Spierdijk. 2012. Assessing competition with the Panzar-Rosse model: The role of scale, costs, and equilibrium. *Review of Economics and Statistics* 94: 1025–1044.

Blackwell, David, and Drew Winters. 1997. Banking relationships and the effect of monitoring on loan pricing. *Journal of Financial Research* 20: 275–289.

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 Anjan Thakor. 2000. Can relationship banking survive competition? *Journal of Finance* 55: 679–713.

Bos, Jaap, James Kolari, and Ryan van Lamoen. 2013. Competition and innovation: Evidence from financial services. *Journal of Banking and Finance* 37: 1590–1601.

Bossone, Biagio, and Jong-Kun Lee. 2004. In finance, size matters: The "systemic scale economies" hypothesis. *IMF Staff Papers* 51: 19–46.

Botsch, Matthew, and Victoria Vanasco. 2015. Relationship lending: Do banks learn? Unpublished manuscript, Bowdoin College and Stanford University, June 17.

Bradley, Michael D., and Dennis Jansen. 1986. On deposit market deregulation and interest rates. *Southern Economic Journal* 53: 478–489.

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.

Brissimis, Sophocles, and Manthos Delis. 2009. Bank-level estimates of market power. Bank of Greece and University of Ioannina, January.

Broecker, Thorsten. 1990. Creditworthiness tests and interbank competition. *Econometrica* 58: 429–452.

Burgstaller, Johann. 2005. Interest rate pass-through estimates from vector autoregressive models. Unpublished Manuscript, Johannes Kepler University.

Calomiris, Charles, and Thanavut Pornrojnangkool. 2009. Relationship banking and the pricing of financial services. *Journal of Financial Services Research* 35: 189–224.

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.

Carbo-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.

Carletti, Elena. 2004. The structure of bank relationships, endogenous monitoring, and loan rates. *Journal of Financial Intermediation* 13: 58–86.

Carlson, Mark, and Kris James Mitchener. 2006. Branch banking, bank competition, and financial stability. *Journal of Money, Credit, and Banking* 38: 1293–1328.

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.

Cenni, Stefano, Stefano Monferrà, Valentina Salotti, Marco Sangiorgi, and Giuseppe Torluccio. 2015. Credit rationing and relationship lending. Does firm size matter? *Journal of Banking and Finance* 53: 249–265.

Cerasi, Vittoria, Barbara Chizzolini, and Marc Ivaldi. 2002. Branching and competition in the European banking industry. *Applied Economics* 34: 2213–2225.

Chang, Chun, Guanmin Liao, Xiaoyun Yu, and Ni Zheng. 2014. Information from relationship lending: Evidence from loan defaults in China. *Journal of Money, Credit, and Banking* 46: 1225–1257.

Chen, Yehnin, Rachel Huang, John Tsai, and Larry Tzeng. 2015. Soft information and small business lending. *Journal of Financial Services Research* 47: 115–133.

Claessens, Stijn. 2009. Competition in the financial sector: Overview of competition policies. International Monetary Fund Working Paper WP/09/05, March.

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.

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.

Coccorese, Paolo. 2009. Market power in local banking monopolies. *Journal of Banking and Finance* 33: 1196–1210.

Cohen, Andrew, and Michael Mazzeo. 2010. Investment strategies and market structure: An empirical analysis of bank branching decisions. *Journal of Financial Services Research* 38: 1–21.

Cornée, Simon, David Masclet, and Gervais Thenet. 2012. Credit relationships: Evidence from experiments with real bankers. *Journal of Money, Credit, and Banking* 44: 957–980.

Corvoisier, Sandrine, and Reint Gropp. 2002. Bank concentration and retail interest rates. *Journal of Banking and Finance* 26: 2155–2189.

Cottarelli, Carlo, and Angeliki Kourelis. 1994. Financial structure, bank lending rates, and the transmission mechanism of monetary policy. *IMF Staff Papers* 41: 587–623.

Cotugno, Matteo, Stefano Monferrà, and Gabriele Sampagnaro. 2013. Relationship lending, hierarchical distance, and credit tightening. Evidence from the financial crisis. *Journal of Banking and Finance* 37: 1372–1385.

Craig, Ben, and Valeriya Dinger. 2010. A microeconometric investigation into bank interest rate rigidity. Federal Reserve Bank of Cleveland Working Paper 10-10, March.

Craig, Steven, and Pauline Hardee. 2007. The impact of bank consolidation on small business credit availability. *Journal of Banking and Finance* 31: 1237–1263.

Cyree, Ken, and W. Paul Spurlin. 2012. The effects of big-bank presence on the profit efficiency of small banks in rural markets. *Journal of Banking and Finance* 36: 2593–2603.

Dai, Mian, and Yuan Yuan. 2013. Product differentiation and efficiencies in the retail banking industry. *Journal of Banking and Finance* 37: 4907–4919.

Davies, Richard, and Belinda Tracey. 2014. Too big to be efficient? The impact of implicit subsidies on estimates of scale economies for banks. *Journal of Money, Credit, and Banking* 46: 219–253.

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 2005.

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. 2010. Bank involvement with SMEs: Beyond relationship lending. *Journal of Banking and Finance* 34: 2280–2293.

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.

Degryse, Hans, and Steven Ongena. 2005. Distance, lending relationships, and competition. *Journal of Finance* 60: 231–266.

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. 2008. Competition and regulation in the banking sector: A review of the empirical evidence on the sources of bank rents. In *Handbook of Financial Intermediation and Banking*, ed. Anjan Thakor and Arnoud Boot, 483–554. Amsterdam: Elsevier.

Delis, Manthos. 2012. Bank competition, financial reform, and institutions: The importance of being developed. *Journal of Development Economics* 97: 450–465.

Delis, Manthos, and Efthymios Tsionas. 2009. The joint estimation of bank-level market power and efficiency. *Journal of Banking and Finance* 33: 1842–1859.
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.
Demsetz, Harold. 1973. Industry structure, market rivalry, and public policy. *Journal of Law and Economics* 16: 1–9.
Devaney, Michael, and William Weber. 2002. Small-business lending and profit efficiency in commercial banking. *Journal of Financial Services Research* 22: 225–246.
DeYoung, Robert. 1998. Management quality and X-inefficiency in national banks. *Journal of Financial Services Research* 13: 5–22.
DeYoung, Robert. 2003. De novo bank exit. *Journal of Money, Credit, and Banking* 35: 711–728.
DeYoung, Robert. 2008. Safety, soundness, and the evolution of the U.S. banking industry. In *Handbook of Financial Intermediation and Banking*, ed. Anjan Thakor and Arnoud Boot, 347–373. Amsterdam: Elsevier.
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, 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, W. Scott Frame, Dennis Glennon, and Peter Nigro. 2011. The information revolution and small business lending: The missing evidence. *Journal of Financial Services Research* 39: 19–33.
Dia, Enzo. 2013. How do banks respond to shocks? A dynamic model of deposit-taking institutions. *Journal of Banking and Finance* 37: 3623–3638.
Dia, Enzo, and Massimo Giuliodori. 2012. Portfolio separation and the dynamics of bank interest rates. *Scottish Journal of Political Economy* 59: 28–46.
Dick, Astrid. 2006. Nationwide branching and its impact on market structure, quality, and bank performance. *Journal of Business* 79: 567–592.
Dick, Astrid. 2007. Market size, service quality, and competition in banking. *Journal of Money, Credit, and Banking* 39: 49–81.
Dinç, L. Serdar. 2000. Bank reputation, bank commitment, and the effects of competition in credit markets. *Review of Financial Studies* 13: 781–812.
Duygun, Meryem, Vania Sena, and Mohamed Shaban. 2013. Shumpeterian competition and efficiency among commercial banks. *Journal of Banking and Finance* 37: 5176–5185.
Edwards, Franklin. 1973. Advertising and competition in banking. *Antitrust Bulletin* 18: 23–31.
Edwards, Franklin. 1976. More on advertising and competition in banking. *Antitrust Bulletin* 21: 85–89.
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.
Elsas, Ralf. 2005. Empirical determinants of relationship lending. *Journal of Financial Intermediation* 14: 32–57.
Elyasiani, Elyas, and Lawrence Goldberg. 2004. Relationship lending: A survey of the literature. *Journal of Economics and Business* 56: 315–330.
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.
Ergungor, O. Emre. 2005. The profitability of bank-borrower relationships. *Journal of Financial Intermediation* 14: 485–512.
Espinoza-Vega, Marco, and Allesandro Rebucci. 2003. Retail bank interest rate pass-through: Is Chile atypical? IMF Working Paper WP/03/112.

Felici, Roberto, and Marcello Pagnini. 2008. Distance, bank heterogeneity, and entry in local banking markets. *Journal of Industrial Economics* 56: 500–534.

Feng, Guohua, and Apostolos Serletis. 2010. Efficiency, technical change, and returns to scale in large U.S. banks: Panel data evidence from an output distance function satisfying theoretical regularity. *Journal of Banking and Finance* 34: 127–138.

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.

Fiordelisi, Franco, Stefano Monferrà, and Gabriele Sampagnaro. 2014. Relationship lending and credit quality. *Journal of Financial Services Research* 46: 295–315.

Fredriksson, Antti, and Andrea Moro. 2014. Bank-SMEs relationships and banks' risk-adjusted profitability. *Journal of Banking and Finance* 41: 67–77.

Fuertes, Ana-Maria, and Shelagh Heffernan. 2009. Interest rate transmission in the U.K.: A comparative analysis across financial firms and products. *International Journal of Finance and Economics* 14: 45–63.

Fuertes, Ana-Maria, Shelagh Heffernan, and Elena Kalotychou. 2010. How do U.K. banks react to changing central bank rates? *Journal of Financial Services Research* 37: 99–130.

Fungàčovà, Laura Solanko, and Laurent Weill. 2014. Does competition influence the bank lending channel in the euro area? *Journal of Banking and Finance* 49: 356–366.

Gambacorta, Leonardo, and Paolo Emilio Mistrulli. 2014. Bank heterogeneity and interest rate setting: What lessons have we learned since Lehman Brothers? *Journal of Money, Credit, and Banking* 46: 753–778.

Gilbert, R. Alton. 1984. Bank market structure and competition: A survey. *Journal of Money, Credit, and Banking* 16: 617–644.

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.

Gilligan, Thomas, William Marshall, and Michael Smirlock. 1984. Scale and scope economies in the multiproduct banking firm. *Journal of Monetary Economics* 13: 393–405.

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, and John Wilson. 2009. Competition in banking: A disequilibrium approach. *Journal of Banking and Finance* 33: 2282–2292.

Goddard, John, Phil Molyneux, and John Wilson. 2004. Dynamics of growth and profitability in banking. *Journal of Money, Credit, and Banking* 36: 1069–1090.

Goddard, John, Hong Liu, Philip Molyneux, and John Wilson. 2011. The persistence of bank profit. *Journal of Banking and Finance* 35: 2881–2890.

Gonzalez, Francisgo. 2009. Determinants of bank-market structure: Efficiency and political economy variables. *Journal of Money, Credit, and Banking* 41: 736–754.

Gual, Jordi. 1999. Deregulation, integration, and market structure in European banking. *Journal of the Japanese and International Economics* 13: 372–396.

Guevara, De, Juan Fernández, Joaquín Maudos, and Francisco Pérez. 2005. Market power in European banking sectors. *Journal of Financial Services Research* 27: 109–137.

Hainz, Christa, Laurent Weill, and Christophe Godlewski. 2013. Bank competition and collateral: Theory and evidence. *Journal of Financial Services Research* 44: 131–148.

Hannan, Timothy. 1991. Foundations of the structure-conduct-performance paradigm in banking. *Journal of Money, Credit, and Banking* 23: 68–84.

Hannan, Timothy. 2006. Retail deposit fees and multimarket banking. *Journal of Banking and Finance* 30: 2561–2578.

Hannan, Timothy, and Robert Adams. 2011. Consumer switching costs and firm pricing: Evidence from bank pricing of deposit accounts. *Journal of Industrial Economics* 59: 296–320.

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 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 Robin Prager. 2004. The competitive implications of multimarket bank branching. *Journal of Banking and Finance* 28: 1889–1914.
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.
Harimohan, Rashmi, Michael McLeay, and Garry Young. 2016. Pass-through of bank funding costs to lending and deposit rates: Lessons from the financial crisis. Bank of England Working Paper No. 590, April.
Hasan, Iftekhar, and Stephen Smith. 1997. A note on fixed costs and the profitability of depository institutions. *Managerial and Decision Economics* 18: 47–53.
Hasan, Iftekhar, William Curt Hunter, and Roswell Mathis III. 2002. Promotional expenditures, market competition, and thrift behavior. *Journal of Business Research* 50: 177–184.
Hattori, Masuzumi, Kohei Shintani, and Hirofumi Uchida. 2015. The repository of soft information within bank organizations. *Journal of Money, Credit, and Banking* 47: 737–770.
He, Zhaozhao. 2015. Rivalry, market structure, and innovation: The case of mobile banking. *Review of Industrial Organization* 47: 219–242.
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.
Hernàndez-Murillo, Rubén, Gerard Llobet, and Roberto Fuentes. 2010. Strategic online banking adoption. *Journal of Banking and Finance* 34: 1650–1663.
Hirtle, Beverly. 2007. The impact of network size on bank branch performance. *Journal of Banking and Finance* 31: 3782–3805.
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' retail rates. *Economica* 71: 99–123.
Homma, Tetsushi, Yoshiro Tsutsui, and Hirofumi Uchica. 2014. Firm growth and efficiency in the banking industry: A new test of the efficient structure hypothesis. *Journal of Banking and Finance* 40: 143–153.
Hristov, Nikolay, Oliver Hülsewig, and Timo Wollmershäuser. 2014. The interest rate pass-through in the Euro area during the global financial crisis. *Journal of Banking and Finance* 48: 104–119.
Hsieh, Meng-Fen, and Cien Chiang Lee. 2010. The puzzle between banking competition and profitability can be solved: International evidence from bank-level data. *Journal of Financial Services Research* 38: 135–157.
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, and Loretta Mester. 2013. Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function. *Journal of Financial Intermediation* 22: 550–585.
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, and Eric Mayer. 2009. Bank behavior, incomplete pass-through, and the cost channel of monetary policy transmission. *Economic Modelling* 26: 1310–1327.
Humala-Acuna, Alberto. 2005. Interest rate pass-through and financial crises: Do switching regimes matter? The case of Argentina. *Applied Financial Economics* 15: 77–94.
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.

Jackson, William III. 1997. Market structure and the speed of price adjustment: Evidence of non-monotonicity. *Review of Industrial Organization* 12: 37–57.

James, Christopher. 1987. Some evidence on the uniqueness of bank loans. *Journal of Financial Economics* 19: 217–235.

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.

Kahn, Charles, George Pennacchi, and Ben Sopranzetti. 1999. Bank deposit rate clustering: Theory and empirical evidence. *Journal of Finance* 6: 2185–2214.

Kahn, Charles, George Pennacchi, and Ben Sopranzetti. 2005. Bank consolidation and the dynamics of consumer loan interest rates. *Journal of Business* 78: 99–133.

Kauko, Karlo. 2005. Bank interest rates in a small European economy: Som exploratory macro level analyses using Finnish data. Bank of Finland Research Discussion Paper 9, 2005.

Kim, Moshe, and Bent Vale. 2001. Non-price strategic behavior: The cost of bank branches. *International Journal of Industrial Organization* 19: 1583–1602.

Kirschenmann, Karolin. 2016. Credit rationing in small firm-bank relationships. *Journal of Financial Intermediation* 26: 68–99.

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.

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. 2012. Imperfect competition in bank retail markets, deposit and loan rate dynamics, and incomplete pass through. *Journal of Money, Credit, and Banking* 44: 1185–1205.

Laderman, Elizabeth. 2007. Market power and relationships in small business lending. Federal Reserve Bank of San Francisco Working Paper 2007-07.

Laderman, Elizabeth. 2008. Small business lending and banking competition. Federal Reserve Bank of San Francisco Economic Letter No. 2008-15, May 9.

Lapp, John. 1976. Market structure and advertising in the savings and loan industry. *Review of Economics and Statistics* 58: 202–208.

Leroy, Aurelien. 2014. Competition and the bank lending channel in the Euroone. *Journal of International Financial Markets, Institutions, and Money* 31: 296–314.

Leroy, Aurelien, and Yannick Lucotte. 2015. Structural and cyclical determinants of bank interest rate pass-through in the Eurozone. Narodowy Bank Polski Working Paper No. 198.

Li, Yutao, Anthony Saunders, and Pei Shao. 2015. The monitoring incentive of transactional and relationship lenders: Evidence from the syndicated loan market. *Journal of Money, Credit, and Banking* 47: 702–735.

Liu, Ming-Hua, Dimitris Margaritis, and Alireza Tourani-Rad. 2011. Asymmetric information and price competition in small business lending. *Journal of Banking and Finance* 35: 2189–2196.

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.

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.

Marrouch, Walid, and Rima Turk-Ariss. 2014. Joint market power in banking: Evidence from developing countries. *Journal of International Financial Markets, Institutions, and Money* 31: 253–267.

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.

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 *Handbook of Financial Intermediation and Banking*, ed. Anjan Thakor and Arnoud Boot, 133–162. Amsterdam: Elsevier.

Mirzaei, Ali, and Tomoe Moore. 2014. What are the driving forces of bank competition across different income groups of countries? *Journal of International Financial Markets, Institutions, and Money* 32: 38–71.

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.

Molnár, József. 2008. Market power and merger simulation in retail banking. Bank of Finland Research Discussion Paper #4-2008.

Montgomery, Heather, Kozo Harmaya, and Yuki Takahashi. 2014. Too big to succeed? Banking sector consolidation and efficiency. *Journal of International Financial Markets, Institution, and Money* 32: 86–106.

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.

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.

Panzar, J.C., and J.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.

Petersen, Mitchell, and Raghuram Rajan. 1994. The benefits of lending relationships: Evidence from small business data. *Journal of Finance* 49: 3–37.

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. 2002. The information revolution and small business lending: Does distance still matter. *Journal of Finance* 57: 2533–2570.

Pilloff, Steven. 1999. Does the presence of big banks influence competition in local markets? *Journal of Financial Intermediation Research* 15: 159–177.

Presbitero, Andrea, and Alberto Zazzaro. 2011. Competition and relationship lending: Friends or foes? *Journal of Financial Intermediation* 20: 387–413.

Puri, Maju, Jorg Rocholl, and Sascha Steffen. 2011. On the importance of prior relationships in bank loans to retail customers. European Central Bank Working Paper No. 1395, November.

Raknerud, Arvid, Bjørn Helge Vatne, and Ketil Rakkestad. 2011. How do banks' funding costs affect interest margins? Norges Bank Financial Stability Working Paper 2011-09, July 5.

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.

Rhoades, Stephen. 1980. Monopoly and expense preference behavior: An empirical investigation of the behavioralist hypothesis. *Southern Economic Journal* 47: 419–432.

Rhoades, Stephen. 1982. Welfare loss, redistribution effects, and restriction of output due to monopoly in banking. *Journal of Monetary Economics* 9: 375–387.

Richards, Timothy, Ram Acharya, and Albert Kagan. 2008. Spatial competition and market power in banking. *Journal of Economics and Business* 60: 436–454.
Rosen, Richard. 2007. Bank market conditions and deposit interest rates. *Journal of Banking and Finance* 31: 3862–3884.
Salop, Steven, and David Scheffman. 1983. Raising rivals' costs. *American Economic Review Papers and Proceedings* 73: 267–271.
Salop, Steven, and David Scheffman. 1987. Cost-raising strategies. *Journal of Industrial Economics* 36: 19–34.
Sander, Harald, and Stefanie Kleimeier. 2004. Convergence in eruo-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.
Sander, Harald, and Stefanie Kleimeier. 2006. Convergence of interest rate pass-through in a wider euro zone? *Economic Systems* 30: 405–423.
Scheffman, David, and Richard Higgins. 2003. Twenty years of raising rivals' costs: History, assessment, and future. *George Mason Law Review* 12: 371–387.
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. 1996. Asymmetric adjustment of commercial bank interest rates: Evidence from Malaysia and Singapore. *Journal of International Money and Finance* 15: 485–496.
Scholnick, Barry. 1999. Interest rate asymmetries in long-term loan and deposit markets. *Journal of Financial Services Research* 16: 5–26.
Scott, John. 1978. Non-price competition in banking markets. *Southern Economic Journal* 44: 594–605.
Shaffer, Sherrill. 1989. Competition in the U.S. banking industry. *Economics Letters* 29: 321–323.
Shaffer, Sherrill. 1993. A test of competition in Canadian banking. *Journal of Money, Credit, and Banking* 25: 49–61.
Shaffer, Sherrill. 1998. Functional forms and declining average costs. *Journal of Financial Services Research* 14: 91–115.
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. 2002. Conduct in a banking monopoly. *Review of Industrial Organization* 20: 221–238.
Shaffer, Sherrill. 2004. Patterns of competition in banking. *Journal of Economics and Business* 56: 287–313.
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. 1990. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *Journal of Finance* 49: 1069–1087.
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.
Sheehan, Richard. 2006. Price stickiness in bank deposit rates: An analysis of intra-market relationships. Unpublished Manuscript, University of Notre Dame.
Shiman, Daniel. 2007. The intuition behind Sutton's theory of endogenous sunk costs. Working Paper, Federal Communications Commission.
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.
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.

Strahan, Philip. 2008. Bank structure and lending: What we do know and do not know. In *Handbook of Financial Intermediation and Banking*, ed. Anjan Thakor and Arnoud Boot, 97–132. Amsterdam: Elsevier.

Sutton, John. 1991. *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*. Cambridge, MA: MIT Press.

Tennant, David, and Richard Sutherland. 2014. What types of banks profit most from fees charged? A cross-country examination of bank-specific and country-level determinants. *Journal of Banking and Finance* 49: 178–190.

Tieman, Alexander. 2004. Interest Rate Pass-Through in Romania and Other Central European Economies. IMF Worker Paper WP/04/211.

Toolsema, Linda, Jan-Egbert Sturm, and Jakob de Haan. 2001. Convergence of monetary transmission in EMU: New evidence. CESifo Working Paper No. 465.

Vajanne, Laura. 2009. Inferring market power from retail deposit interest rates in the euro area. Bank of Finland Research Discussion Paper 27/2009.

van der Cruijsen, Carin, and Maaike Diepstraten. 2015. Banking products: You can take them with you, so why don't you? De Nederlandsche Bank Working Paper No. 490, December.

van Leuvensteijn, Michiel, Christoffer Kok Sørensen, Jacob Bikker, and Adrian van Rixtel. 2013. Impact of bank competition on the interest rate pass-through in the euro area. *Applied Economics* 45: 1359–1380.

VanHoose, David. 1984. Bank market structure and monetary policy. Unpublished Ph.D. Dissertation, University of North Carolina at Chapel Hill.

VanHoose, David. 1985. Bank market structure and monetary control. *Journal of Money, Credit, and Banking* 17: 298–311.

VanHoose, David. 1988. Deposit market deregulation, implicit deposit rates, and monetary policy. *Atlantic Economic Journal* 16: 11–23.

VanHoose, David. 2013. Implications of shifting retail market shares for loan monitoring in a dominant-bank model. *Scottish Journal of Political Economy* 60: 291–316.

Von Rheinbaben, Joachim, and Martin Ruckes. 2004. The number and closeness of bank relationships. *Journal of Banking and Finance* 28: 1597–1615.

Weill, Laurent. 2008. Convergence in banking efficiency across European countries. Manuscript, University of Robert Schuman, January.

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.

Wheelock, David, and Paul Wilson. 2012. Do large banks have lower costs? New estimates of returns to scale for U.S. banks. *Journal of Money, Credit, and Banking* 44: 172–199.

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.

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.

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: Wiley.

Xu, Bing, Honglin Wang, and Adrian van Rixtel. 2015. Do banks extract informational rents through collateral? Bank for International Settlements Monetary and Economic Department Working Paper 522, October.

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.

Chapter 4
The Economics of International Banking

As discussed by Lewis and Davis (1987), international banking broadly refers to the "cross-border and cross-currency facets of banking business." This definition accords with a survey of international banking by Aliber (1984), which referred to "[t]wo distinct sets of issues. . .involved in the analysis of international banking." One set of issues is international finance concerns "involv[ing]. . .the role of banks in cross-border and cross-currency financial flows." The other set entails industrial organization issues "center[ing]. . .on the patterns of expansion of foreign branches and subsidiaries of banks." This activity, called multinational banking, involves foreign direct investment in facilities owned and directly operated in host nations separate from the enterprise's home office.

The issue of determinants of net cross-border flows of banking services has taken on a renewed urgency in the wake of the international financial meltdown of 2007–2009 and continuing pressures on global banking markets that since have emerged as a consequence of the Euro area's sovereign debt weaknesses. A considerable portion of the recent literature has examined such issues in the context of purely international-finance perspectives relating to management of globally diversified asset portfolios. Part of this literature is touched upon below, although much of it lies largely outside the array of topics relevant to this book—see Goldberg (2009). Nevertheless, in this chapter you will learn that a considerable amount of recent literature in banking has sought to explore these policy-relevant issues by extending industrial-organization-related analyses to the international realm. One outcome has been an expansion in the study of configurations of multinational banking—that is, either so-called "greenfield entry" via establishment of foreign bank branches in host nations or acquisitions of previously existing host-country banks—emphasized by Aliber in his earlier review. Another consequence, however, has been the extension of real-resource-based bank models to consideration of broader determinants of cross-border flow of banking services. Let's begin, however, by reviewing briefly the two primary waves of growth of global banking.

© Springer-Verlag GmbH Germany 2017

D. VanHoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-662-54326-9_4

The Growth of Global Banking

Most scholars, such as Curry et al. (2003), identify two key periods of growth in international banking. The first of these accompanied European colonialism and U.S. expansionism stretching across much of the nineteenth century but particularly during the interval between 1870 and 1914. The second period of banking globalization has stretched over roughly the past 50 years but especially the period since the early 1990s.

The First Big Wave of Banking Globalization: The Colonial Period

Curry et al. (2003) provide a helpful survey of the nineteenth-century wave of growth in international banking. This period of rapid growth of global banking naturally began with British multinational banks that expanded their operations abroad with the spread of the British empire's mercantile and associated financial interests in colonial holdings in Australia, the Caribbean, and North America in the 1830s and in Asia, Latin America, and Southern Africa in the 1850s.

British banks additionally expanded their multinational operations in the Middle East following the 1870s, and after this date banks of other European countries, such as Belgian, France, Germany, and the Netherlands embarked on global-banking buildups as these nations increased the scope of their colonial pursuits. By the 1890s, U.S. banks joined in with the advent of U.S. international expansionism.

The Second Globalization Wave: The Most Recent Decades

Growth of international banking was halted and then persistently slowed by the outbreak of World War I, European financial-market chaos during the interwar period, renewed breakdown of international financial relationships during World War II, and protectionist financial policies—in particular, restrictions on cross-border financial-capital flows—enacted during the post-war recovery. Curry et al. (2003), following Huertas (1990)—also see Minoiu and Reyes (2013) for a network-based analysis—point to three fundamental factors that eventually fueled a global-banking resurgence. One was a macroeconomic boom that took place as the global recovery reached a full swing by the 1960s. A second factor was an effort by a number of multinational banks—especially in the United States in the face of deposit-interest-rate ceilings—to adapt to regulations that artificially constrained quantities supplied of domestic deposit funds by seeking those funds abroad. A third factor was technical change—in particular, advances in telecommunications

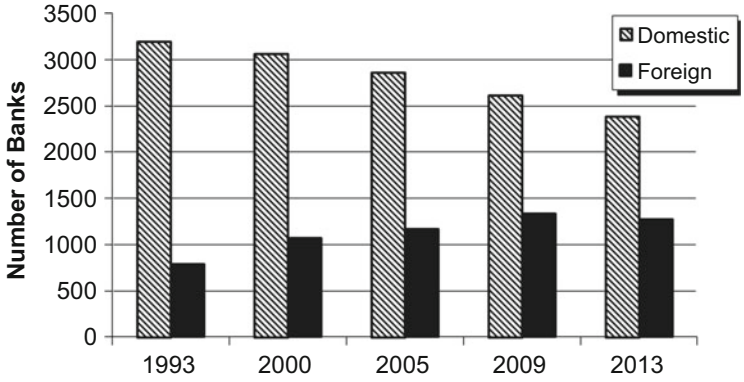

Fig. 4.1 Numbers of domestic and foreign banks in more than 120 host countries (Source: Claessens and van Horen 2015, http://www.dnb.nl/en/onderzoek-2/databases/index.jsp)

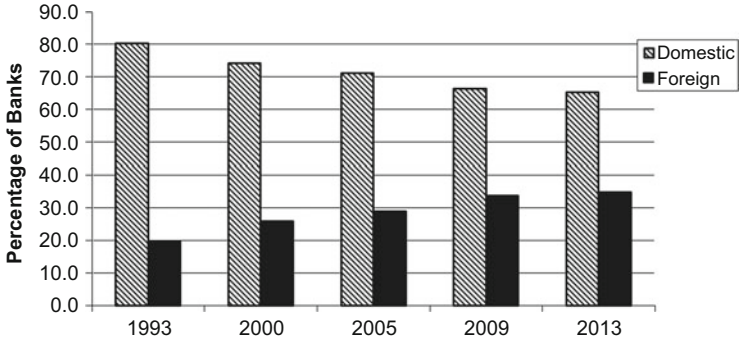

Fig. 4.2 Percentage shares of numbers of domestic and foreign banks in more than 120 host countries (Source: Claessens and van Horen 2015, http://www.dnb.nl/en/onderzoek-2/databases/index.jsp)

and then computing technologies—that reduced the costs of engaging in cross-border banking.

By the late 1980s, many banks based throughout the developed world and even a number based in developing and emerging nations were involved in international-banking activities. Claessens and van Horen, (2012, 2014a, b, 2015) have compiled a continuously updated dataset for most of the subsequent interval. Based on data from 5324 banks in 137 nations, Claessens and van Horen (2014b) document that even as the number of domestic banks within their dataset declined by 18% between 1995 and 2009, the host-country presence of foreign-owned banks rose by 69%. As shown in Fig. 4.1, the net result was a significant rise in the number of foreign banks in host countries by 2009.

Inclusive of updated data that takes into account the years encompassing the global financial crisis of 2007–2009, Claessens and van Horen (2015) determine that a slight drop subsequently took place in the number of foreign banks operating worldwide following the onset of the crisis. Nevertheless, as indicated in Fig. 4.2,

their data reveal a continuing increase in the percentage share of foreign banks. Overall, however, they find a small dip in average foreign-bank asset share since the crisis, from about 13% to about 11%. Adams-Kane et al. (2015) investigate the channels through which the provision of cross-border loans were affected by the crisis and find that solvency problems caused by weakened balance sheets and, in particular, altered willingness to lend because of greater economic uncertainty and reduced interbank-funding availability adversely affected flows of credit during the crisis. Bremus and Fratzscher (2014) find that toughened capital regulation in response to the crisis generated regulatory-arbitrage spillover effects that reduced cross-border bank lending, although monetary policy actions separately helped somewhat to regenerate banks' cross-border lending activities.

Fundamental Elements Influencing Cross-Border Banking

Much of the literature on international banking seeks to explore empirically the determinants of the extent of banking globalization based on appeals to likely incentives to engage in cross-border banking activities. One branch of the literature [see, for instance, Beck (2002), Battilossi (2006), Do and Levchenko (2007), and Jaud et al. (2014)] has focused primarily on trade flows in markets for goods and services as causal variables in the development of global banking. With respect to the elements that influenced the initial banking globalization wave between 1870 and the First World War, Battilossi (2006) uses the number of foreign branches of multinational British, French, and German banks as the dependent variable and considers a number of potential explanatory variables, including geographical factors, international trade and capital flows, institutional characteristics, and a variety of economic features of host nations. He finds no evidence that any specific variables stand out as determinants of this wave of global banking. Battilossi concludes that institutional and economic elements played heterogeneous roles in shaping the globalization wave, with various combinations of the independent variables helping to best explain banking globalization for different nations over this interval.

Distance

Many empirical studies include some of the same basic explanatory variables, however. Consistent with empirical "gravity" specifications commonly employed in the international trade literature, the primary geographical factor examined in studies of determinants of international banking activities is distance. Naturally, a widely used distance variable is the physical separation, measured in kilometers or miles, between the nation in which an individual bank is based and a country in

which it engages in cross-border banking activities. In addition, however, researchers commonly employ cultural-distance measures, such as common languages or language roots. For instance, Buch (2003) find evidence of effects of such distance measures on foreign asset holdings of banks based in eight nations with activities in 75 host countries between 1983 and 1999.

As Claessens and van Horen (2014a) note, in theory information costs related to distance measures can exert contrary incentives on decisions regarding the scope of international banking activities. On one hand, increased distance from customers can adversely influence banks' capabilities to collect and transmit "soft" information. On the other hand, increased distance from competitors may increase the likelihood that a bank earns higher net returns from its international banking activities. Claessens and van Horen provide evidence, using bilateral data on 1199 foreign banks from 75 home nations that provide services in 110 host countries, that both bilateral distance from customers and the weighted average distance of all competing banks to a host country—"competitive remoteness"—affect the a foreign bank's decision about involvement in a host-country market.

Kleimeier et al. (2013) also emphasize the potential role of distance effects. They utilize BIS Locational Banking Statistics in order to examine quarterly 1995–2008 data for 23 nations in which banks are based and 165 countries in which customers are located. The authors, who focus mainly on impacts of currency crises on cross-border banking flows, employ a gravity model in which the relevant size variable is the product of nations' GDP levels and in which distance between nations proxies for transaction costs. They find that distance effects are stronger determinants of cross-country deposit flows than flows of loans.

As discussed in greater detail below, the extent of the international scope of banks' activities hinges in large part on how a broadening of global scope affects their performances. Claessens and van Horen (2012) analyze, using an earlier version of their dataset, the profitability performances of foreign banks relative to domestic banks between 1999 and 2006. Claessens and van Horen find little evidence that the relative performance of foreign banks is improved by closer physical proximity. They find that foreign banks based in high-income nations and operating in weakly regulated host nations tend to exhibit stronger relative performances. Closer language proximity and similar regulatory structures also contribute to better relative performances of foreign banks.

Competitive Structure and Effects of Entry in Host-Country Banking Markets

Some research has explored the interaction among host-banking-market concentration, bank size, and international banking. One issue relates to terms of lending on the part of domestic versus foreign banks. Beck et al. (2011) analyze data regarding SME lending terms from a 2007 survey of 91 banks in 45 nations. They find that foreign banks are more likely to make collateralized loans, to make use of soft information in lending, and to utilize centralized loan-approval processes.

Otherwise, the Beck et al. conclude that there are few distinctions between domestic and foreign banks in terms of lending. The main differences arise in comparing loan terms in developed and developing nations as a consequence of differences in institutional and legal settings.

Regarding foreign-competition effects, Müller and Uhde (2013) employ 1993–2007 BIS Consolidated Banking Statistics data to investigate determinants of the cross-border lending of banks in 13 OECD nations. They find evidence that banks already operating in concentrated home OECD markets tend to be less likely to engage in international banking in emerging-economy host markets than banks that confront more intense domestic competition. They also find that banks with lower levels of capitalization are more likely to lend in emerging nations. In a study of expansion by Kenyan banks into other East African banks between 2002 and 2012, Kodongo et al. (2015) find some evidence indicating that intense competition in the home market also can provide incentives for domestic banks to look for profitable opportunities abroad.

Of course, cross-border banking influences the degree of competition in a host nation's banking markets. Several studies have focused on this issue. For instance, Jeon et al. (2011) employ banking data for 17 Asian and Latin American countries over the 1997–2008 interval to assess the effects of foreign bank presence in host nations on the Panzar-Rosse measure of competition developed by Bikker and Haaf (2002) and conclude that foreign penetration of host nations' banking markets unambiguously boosts competition, both directly and via spillover effects onto host nations' domestic banks. In addition, Manlagñit (2011) evaluates the effects of liberalization of foreign-bank entry barriers on the performances of 31 domestic Philippine banks using data over the 1990–2006 interval. Manlagñit considers how foreign bank presence in the Philippines—measured by the foreign share of either the number of banks or of total bank assets—affect bank profitability and cost efficiency, and she finds that domestic banks' profit rates and costs both declined as foreign bank penetration increased, consistent with pro-competitive effects of foreign bank entry.

In contrast to pro-competitive effects of cross-border banking found in the preceding studies, Gormley's (2010) assessment of the effects of foreign-bank entries into Indian markets during the 1990s indicates that foreign banks financed only a small array of highly profitable firms upon entry and that foreign bank entry crowded out about 8% points of domestic lending. The crowding-out effect on domestic credit fell most heavily on smaller firms dependent on external financing and adversely affected their performance. It also generated a larger decline in home lending to firms with fewer tangible assets and firms affiliated with business groups, which Gormley argues is consistent with worsened informational asymmetries as a consequence of foreign bank entries.

To the extent that international banking does engender more intense competition, it is likely to push down retail loan rates and boost retail deposit rates. Popov and Ongena (2011) examine data on loan contract terms for more than 6000 firms in 14 European nations between January 1998 and December 2005. These authors find evidence that increased interbank market integration across these European

countries during this period was associated with reduced stringency of borrowing constraints and significantly lower loan interest rates, particularly in nations with the most competitive and integrated markets. In a separate analysis of European data, Rughoo and Sarantis (2014) consider nine sets of monthly data for loan and deposit rates applicable to households in 15 European Union countries over the 2003–2011 interval and evaluate the degree of convergence through the end of 2007 and after the beginning of 2008. Rughoo and Sarantis find evidence of convergence of loan rates and, in particular, deposit rates across the EU markets until the outset of the financial crisis, after which rates of convergence either slowed considerably or even diverged.

A contrary analysis provided by Clerides et al. (2015), who estimate degrees of competition in 148 nations during the 1997–2010 interval, indicates more intense competition may not have accompanied the recent wave of cross-border banking. These authors find that overall, the intensity of banking competition across these nations deteriorated between 1997 and 2006, briefly increased until 2008, and then began to deteriorate once more.

Bank Size and Entry Costs

Relative sizes of competing banks and entry costs confronted by prospective entrants likely also matter. Bremus (2015) contemplates the size issue by extending a theoretical model of endogenous bank markups developed by de Blas and Russ (2010) to analyze effects on cross-border banking on concentration in retail markets containing banks of differing sizes and lending efficiencies. Banks engage in Bertrand competition by undercutting lending rates of rivals within niches until the most efficient bank captures total credit demand within each niche, so that a range of loan-rate markups emerges as an equilibrium outcome. Bremus extends this underlying structure to a two-country setting and conducts simulations using calibrated values of parameters in order to obtain three empirical predictions: (1) cross-border lending generates lower concentration in domestic credit markets; (2) increased banking-sector foreign direct investment increases the overall efficiency of lending; and (3) efficiency gains boost banks' net interest margins even as concentration decreases.

Several studies have explored empirical effects of bank size and entry costs on the scope of international banking. Buch et al. (2011) seek to evaluate whether firm size and productivity heterogeneities, which have been found to influence international activities of manufacturing firms, help to explain the foreign involvement of banks. Based on a set of data for 2226 German banks and their activities in 63 nations between 2002 and 2006, the authors conclude that while the bulk of international banking is conducted by the largest banks, most banks and even a large share of the smallest banks had a presence abroad. Nevertheless, larger banks active in more distant markets and hold substantially greater volumes of assets abroad. Banking productivity and size are not perfectly correlated, however. Although there is a positive relationship between banks' productivities and the

extent of their foreign activities, the authors find that a number of highly productive banks have little involvement in foreign markets and that some of the least productive banks have significant asset holdings abroad in relation to their size.

Temesvary (2015) studies quarterly data for the 1997–2013 period in 107 banking markets of foreign hosts of U.S. international banking activities. Consistent with Buch et al.'s findings regarding German banks, Temesvary finds no evidence of a monotonic relationship between bank size and the breadth of foreign involvement. Small banks have boosted foreign activities in relation to growth of their assets at a faster pace than larger banks. Temesvary finds that the scope of a bank's foreign involvement is positively related to prior experience in other countries. Separate estimations of determinants of cross-border versus foreign-affiliate claims of U.S. banks indicates that banks regard the two mechanisms of international banking as substitutes with weaker host regulation and stronger host economic conditions leading to greater use of foreign affiliates. Fixed foreign entry costs confronted by banks are positively related to general costs of doing business in host nations and to the stringency of regulations in those countries. Within these data, mean foreign entry costs faced by U.S. banks for the entire period are about $78 million.

In a separate study, Temesvary (2014) additionally examines 2003–2013 quarterly data for 82 U.S. banks with foreign operations in 83 nations. She finds evidence that bank characteristics are stronger determinants of foreign activities than host-country traits. Furthermore, Temesvary concludes that banks—particularly more highly capitalized banks—have shifted toward a greater reliance on foreign subsidiaries relative to direct cross-border activities, that the global financial crisis induced both increased risk aversion on the part of banks and their regulators and more than a tripling of the fixed costs of bank entry into foreign markets, from a pre-crisis mean of about $71 million, close to the value obtained in her other study.

Idiosyncratic Elements and Prior Experience

Researchers have contemplated a range of other factors that can affect the scope of banks' international activities. Shirota (2015) explores the role of idiosyncratic factors by applying a Bayesian dynamic latent factor model to decompose volatilities of cross-border credit flows across 70 countries to assess global-, regional-, and country-specific-common-factor contributions to those flows. The author, who utilized BIS data on banks' external claims vis-à-vis banking and non-banking sectors, finds that overall, global and regional common factors account for slightly less than half of these volatilities, but significant heterogeneities exist across countries.

Undoubtedly, learning-by-doing effects can exert an important influence on banks' cross-border activities. Clare et al. (2013) focus on elements that affect banks' choices to continue foreign operations once those operations already have been established. The authors utilize data on 408 London operations of 77 banks

based in nations outside the United Kingdom during the 1945–1999 interval. They find that the likelihood of continuation of the banks' London activities is positively related to the level of global economic activity and negatively related to the volatility of global economic activity. In addition, continuation of banks' international operations in London depends positively on banks' level of prior experience with those activities, the size of their operations, and the extent to which their activities are integrated with the local marketplace upon initial entry.

The Industrial Organization of Multinational Banking

Considerable research in recent years has contemplated traditional industrial organization issues related to foreign direct investment on the part of multinational banks. Part of this work has examined determinants of the overall scope of multinational activities. Other research has examined distinctive effects of multinational banks' choice between investing by opening new offices in host countries—greenfield investment—versus investment undertaken via acquisition of previously existing host-country institutions.

Determinants of the Overall Scope of Foreign Direct Investment in Banking

What key elements determine the level of foreign direct investment (FDI) undertaken by multinational banks? Berger et al. (2000) argue that banks seeking to expand into host nations can utilize advantages possessed in their home countries to overcome possible efficiency disadvantages of multinational operations, although Sturm and Williams (2010) find little evidence to support this hypothesis in Australian data.

Recent research emphasizes the role of implicit internal capital markets operating within multinational institutions. Dietrich and Vollmer (2010) focus on relative advantages of engaging in cross-border lending from home offices in nations in which banks are based versus engaging in multinational banking by establishing or acquiring subsidiaries in other countries. These researchers argue that that the decision hinges on a key benefit of multinational banking—the ability to create more liquidity—versus a fundamental cost—resulting inefficiencies in internal (within-bank) capital markets on which banks rely to allocate liquidity across countries. Dietrich and Vollmer also suggest that toughened capital regulation contributes to less reliance on internal-capital-markets funds transfers and consequently give banks greater incentives to become multinationals instead of engaging in cross-border lending.

García-Herrero and Vázquez (2013) evaluate banks' gains from international diversification by studying data on 38 multinational banks and their foreign subsidiaries between 1995 and 2004. They find evidence that multinational banks' risk-adjusted returns are positively related to shares of assets allocated to subsidiaries' operations in emerging economies and more geographically widespread dispersions of subsidiaries across differing regions.

De Haas and van Lelyveld (2010) emphasize the importance of implicit internal capital markets operating within multinational banks and provide evidence that such markets operate within 45 of the largest multinational banks. De Haas and van Lelyveld find evidence that foreign subsidiaries of the financially strongest multinational banks exhibit fastest credit growth and are able to continue to provide credit during financial crisis episodes. Presumably these robust institutions are best positioned to effectively maintain and efficiently operate internal capital-market-style relationships with affiliates.

Finally, analysis of cross-border data from 25 nations' banking systems on the part of Reinhardt and Riddiough (2015) provides evidence that international flows of intragroup, internal-capital-markets-related flows are stable sources of funds. Even during the recent crisis period, these flows were unrelated to both global and local macroeconomic shocks, in contrast to much more volatile interbank lending.

In an analysis of quarterly data from U.S. branches of foreign banks between 2010 and 2012, Cetorelli and Goldberg (2016) find evidence that the organizational complexity involved in structuring large global banks alters the sensitivity of their adjustments of lending to variations in funding costs. Cetorelli and Goldberg's results indicate that maintaining their complex operational scales induces global banks to be about 30% less sensitive to changes in funding costs—in their study, a significant increase in deposit insurance premiums—than other banks.

Greenfield Investment Versus Acquisitions of Host Nations' Banks

Does it matter whether banks initiate foreign operations through trade from its own domestic offices or via foreign direct investment in offices in host nations? If banks choose the latter approach and become multinationals, should they open new subsidiaries or purchase existing host-country institutions? Lehner (2009) conducts a theoretical analysis of trade-offs confronting a bank in considering whether to engage in international banking operations in a host country via cross-border lending versus either greenfield- or acquisition-based foreign direct investment in host-country subsidiaries. Lehner does so using a model in which a foreign bank is assumed to possess no soft information about borrowers and limited information about local banking markets and consequently engages in cross-border lending only if its credit-screening efficiency exceeds a threshold minimum consistent with earning nonnegative profits. Greenfield-based FDI yields better information about

local markets, while acquisition-based FDI yields all soft information previously accumulated by host target banks. The state of a host nation's development influences the ultimate nature of these trade-offs. Lehner concludes that banks will be likely to opt either for cross-border lending or acquisition-based FDI in less developed countries and to acquisition-based FDI entry in developed nations.

Of course, the relative efficiency of greenfield offices versus acquired subsidiaries surely matters for the decision about how to engage in FDI. Mulyaningsih et al. (2015) contemplate the effects of foreign bank's greenfield- and acquisition-based foreign direct investment within Indonesia's banking sector between 1980 and 2000. Mulyaningsih et al. apply the Panzar-Rosse methodology as employed by Bikker and Haaf (2002) and discussed in Chap. 2 and find evidence that foreign banks' subsidiaries, particularly those launched via greenfield FDI, operated at smaller and more efficient scales and thereby behaved more competitively as revealed by lower loan rates and larger relative amounts of lending.

The choice between greenfield- versus acquisition-based entry also may affect market competition and performance in the host nation's retail banking markets. Havrylchyk (2012) examines data for more than 2000 Central and Eastern European nations during the 2000–2005 interval and investigates the effects of the presence of foreign bank subsidiaries on firm entries and exits. Havrylchyk finds sharply contrasting effects of foreign banks' acquisitions, which are associated with reduced firm entry, smaller size of entrants, and increased exits, versus greenfield foreign banks, which boost firm entries as well as exits. She also considers the potential role of informational opaqueness of industries and finds evidence that foreign bank presence has larger effects on entries and exists in more opaque industries.

Correa (2009) contemplates the performance issue by examining data on 220 acquisitions of banks in host nations on the part of multinational banks between 1996 and 2003. Correa finds evidence that banks most likely to be acquired in such deals are large banks in small countries, banks that exhibit weak performances prior to being acquired, and banks that compete in relatively more concentrated local markets. He also concludes that target banks' performances failed to improve relative to other host-country institutions within the first 2 years following acquisition.

Beccalli and Frantz (2009) also consider performance differences. They explore how 714 European-bank merger deals involving acquisitions worldwide between 1991 and 2005 affected the performances of the acquiring banks. Beccalli and Frantz find evidence of slight negative effects on returns on equity and cash flow and on profit efficiency but significant improvement in cost efficiency. They suggest that the cost-efficiency improvements are captured mainly by banks' customers, presumably in the forms of lower loan rates and fees and higher deposit rates.

Degryse et al. (2012) utilize quarterly data from 110 bans in Poland for the 1996–2006 period to evaluate whether the observation of lower foreign-owned bank lending rates and interest spreads exist because of foreign banks are more efficient or because of different allocations of lending relating to borrower

transparency and currency denomination and maturity of these banks' loans. Degryse et al. find that the greenfield entrants typically initially allocate less of their lending to informationally opaque customers than domestic banks and that this fact accounts for observations of lower loan rates and smaller interest spreads. Greenfield foreign banks also extend more loans in foreign currencies and at shorter maturities, though following the passage of time differences with domestic banks regarding opaqueness of lending and currency denomination tend to dissipate.

Real-Resource-Based Analyses of Cross-Border Trade in Banking Services

As the above discussion suggests, in the case of economic analysis of international banking, empirical research has tended to run well ahead of guidance that in-depth theoretical analyses might yield. In contrast to the emphasis that the international-banking literature has placed on motives relating to financial capital allocation via cross-border banking with or without foreign direct investment in host nations, the international trade literature traditionally focuses on incentives relating to comparative advantages arising from firms' utilization of real resources.

In banking as in other industries, such comparative advantages would arise from differences in costs of producing tradable financial services. Although the international-banking literature's longstanding emphasis on financial-capital-allocation incentives seems natural in light of the roles that financial institutions perform, it is also true, as shown in Fig. 4.3, that considerable portions of nations' banks' operating costs relate to expenditures on real resources including labor and physical capital. Indeed, Fig. 4.4 indicates that the average cross-national real resource costs typically account for close to half of all bank expenses. It is not surprising, therefore, that Dia and Menna (2016) have provided evidence that marginal resource costs can explain a large portion of banks' interest rate spreads.

Just a handful of researchers contemplated theories of international banking rooted in standard economic models more typically applied to analysis of international trade in goods and services. Neu (1988), Sagari (1989), Moshirian (1994), Li et al. (2003), Moshirian et al. (2005) are among the very few examples of work that has sought to explore broad implications of existing models of international trade in services for the analysis of financial services trade. Among these studies, that of Sagari (1989) is the sole instance of an effort to develop an explicit international-trade-based model of global trade in financial services. Sagari's analysis, however, is based on a very rudimentary model of financial firms.

Only recently have theories emerged that have contemplated central roles for banking firms as within real-resource-based models of trade. Let's consider each of these theories and the empirical evidence offered to date regarding their potential applications to international banking.

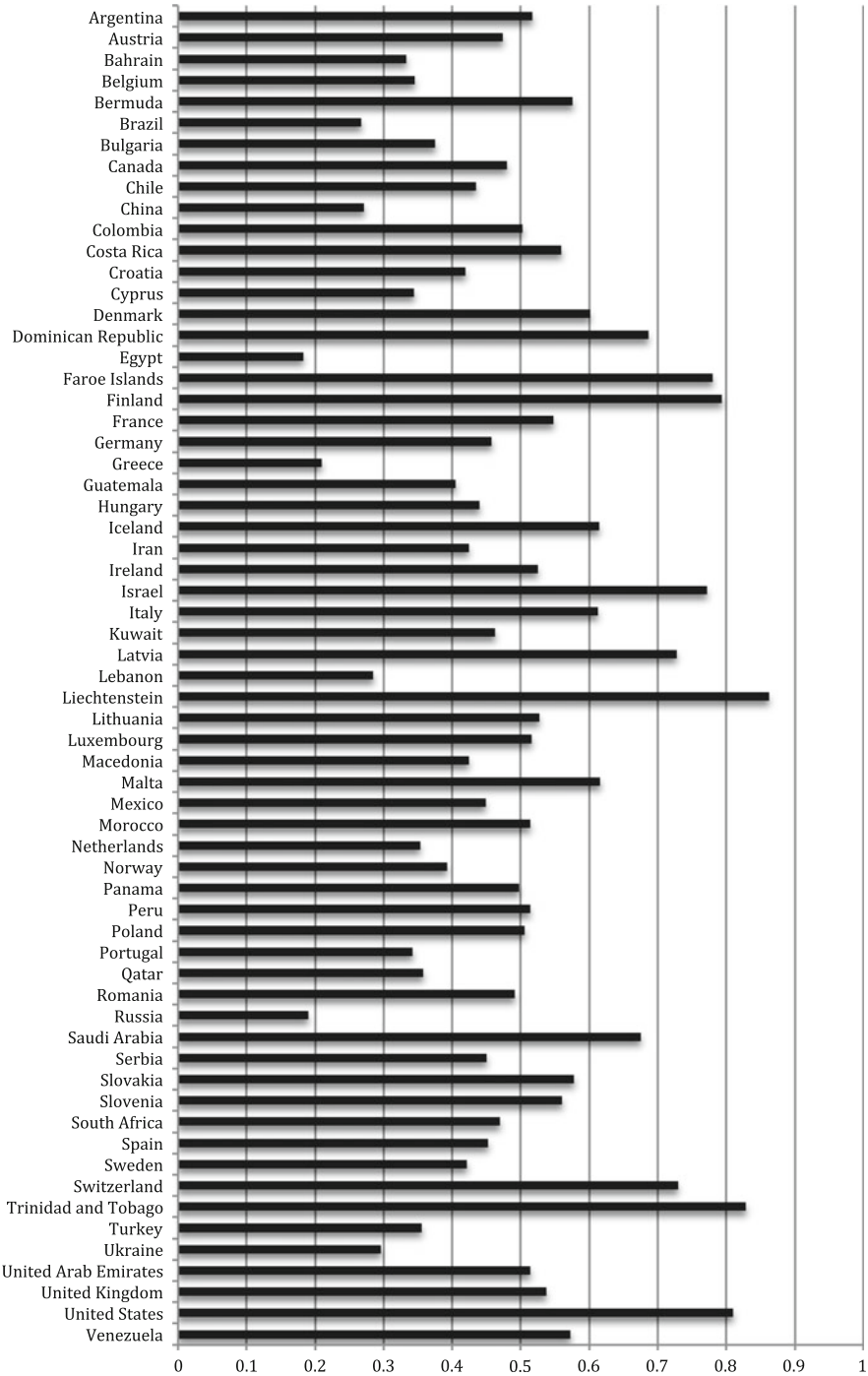

Fig. 4.3 Banks' real resource costs as a share of total expenses, 2015 (Source: Dia and VanHoose 2017)

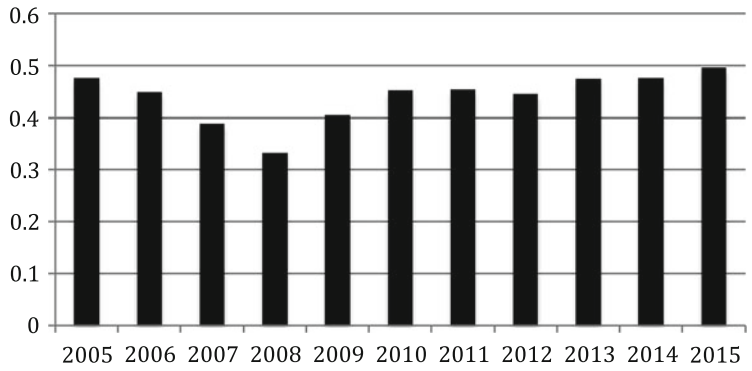

Fig. 4.4 Sixty-three-nation average of bank real resource costs as a share of total expenses (Source: Dia and VanHoose 2017)

Differences in Nation's Capital-Labor Ratios and Efficiencies as Determinants of Cross-Border Banking

Traditional theories of international trade in goods and services focus on national comparative advantages arising from differing relative factor endowments and technological efficiencies. A theoretical framework provided by Niepmann (2015) focuses on these elements to motivate cross-border trade in banking services.

Niepmann considers a setting in which individuals who own physical capital may choose to either to become depositors at banks that lend funds to entrepreneurs or to become entrepreneurs who apply their endowments of physical capital (which they can augment by obtaining bank loans) directly to production. In her model, satisfaction of arbitrage conditions relating financial rates of return to rates of return on real physical capital resources determines cross-border resource flows that can include cross-border banking flows. Niepmann's framework yields a rich set of potential equilibrium possibilities that can encompass what she terms purely *domestic banking* without trade, *international banking* in which banks lend domestic capital abroad, *global banking* in which banks raise foreign capital that they lend abroad, or *foreign sourcing* in which banks borrow abroad to engage in domestic investment.

The Structure of Niepmann's Framework for a Closed Economy

In Niepmann's model, K individuals, called *capitalists*, each own one unit of physical capital, so total physical capital equals K. Those capitalists who opt to hold real deposits with banks place their physical capital in bank deposit accounts that offer a required gross return equal to $1 + r$, where r is the required rate of return in the economy's financial sector. Capitalists who function as entrepreneurs apply a

fixed amount of capital $z > 1$ and variable quantity of labor, denoted here as n, to production of consumption goods with a price normalized at unity utilizing a constant-returns-to-scale technology and production function denoted $F(n, z)$. Because each entrepreneur is endowed with only a single unit of physical capital but $z > 1$ units are required for a productive investment, the entrepreneur must obtain $x = z - 1$ units of physical capital using credit obtained from banks.

Niepmann assumes that banking markets are perfectly competitive and collect a loan service fee c that is proportionate to each loan magnitude x, so in this stark banking environment, therefore, the size of the intermediation-cost parameter c measures the degree of bank efficiency in performing their operations. Firms also operate in perfectly competitive markets. They are assumed to be identical and consequently employ the same quantities of fixed capital, z, and labor, n. The total quantity of available labor is N. If we denote the number of firms as f, then equilibrium in the market for physical capital requires that the number of firms must equal $f = K/z$. Labor-market equilibrium likewise requires that $N = fn$.

Standard marginal-productivity conditions under perfect competition in input markets apply at each firm, so the gross return to physical capital is equal to $R = 1 + F_z(z, n) = 1 + F_z(1, z/n) = 1 + F_K(1, K/N)$, and the real wage rate is equal to $\omega = F_n(z, n) = F_N(1, K/N)$. These conditions thereby imply that the gross return to physical capital and the wage rate depend on the economy's aggregate ratio of physical capital to labor, K/N.

In equilibrium, capitalists are indifferent between acting as depositors or entrepreneurs, so the gross return on deposits, $1 + r$, must equal the net profit earned through entrepreneurial activities, which equals $zR - c(z - 1) - (1 + r)(z - 1)$. This condition yields the following solution for the gross financial return:

$$1 + r = R - c\left(\frac{z-1}{z}\right) = 1 + F_K(K/L) - c\left(\frac{z-1}{z}\right).$$

This condition implies that the rate of return r in the economy's financial sector depends on its endowments of physical capital and labor and the efficiency of its banks. The financial interest rate is lower if the endowment of physical capital is larger in relation to the endowment of labor or if banks are less efficient, so that the intermediation cost is higher.

Extending Niepmann's Setup to a Two-Country Environment

Now suppose that there are two nations, Country 1 and Country 2, that possess potentially different endowments of physical capital and labor and that have unique banking technologies and consequently possibly divergent costs of intermediation. Workers, entrepreneurs, and deposits are assumed to be immobile.

Niepmann assumes that if a bank in Country j lends to firms in Country i, with $i \neq j$, it must incur an additional intermediation cost τ_{ij} that is proportionate to the

amount of the loan. In addition, if the bank borrows abroad, in addition to the interest rate per unit of physical capital that it raises from foreign holdings of real deposits, the bank must pay an additional interest premium t_{ij}. As a consequence, an entrepreneur in Country i confronts the following four options:

(i) Purely domestic banking: pay a domestic bank $c_i + 1 + r_i$ per unit of physical capital that the domestic bank obtains domestically;
(ii) International banking: pay a foreign bank $c_j + \tau_{ij} + 1 + r_j$ per unit of physical capital that the foreign bank obtains from foreign depositors;
(iii) Global banking: pay a foreign bank $c_j + 1 + r_i + t_{ij}$ per unit of physical capital that the foreign bank obtains from domestic depositors; and
(iv) Foreign sourcing: pay a domestic bank $c_i + 1 + r_j + t_{ij}$ per unit of physical capital that the domestic bank obtains from foreign depositors.

Niepmann shows that if the efficiency difference between the two nations' banking systems are not too large, but Country j is endowed with relatively more physical capital compared with labor than Country i, then in equilibrium Country j will be more likely engage in international banking. That is, Country i's domestic entrepreneurs will obtain physical capital from banks in Country j, which in turn will obtain physical capital from Country j's depositors. The resulting flow of physical capital to Country i equalizes the gross return on physical capital across the two nations and thereby maintains global capital-market equilibrium.

In contrast, for a given capital-labor ratio, a sufficiently greater efficiency advantage for Country j's banks generate foreign sourcing in Country j. Entrepreneurs in Country j will tend to borrow physical capital from their domestic banks, which will take advantage of their efficiency advantage to obtain physical capital from Country i's depositors.

In the case in which Country j has both a higher ratio of physical capital to labor than Country i and in which Country j's banks are significantly more efficient than Country i's banks, Country j is more likely to engage in both international banking *and* global banking, with entrepreneurs in both nations obtaining physical capital from Country j's banks, which in turn obtain physical capital from depositors in both nations.

For a situation in which Country j's banks are more efficient but in which that nation is endowed with a relatively lower ratio of physical capital to labor as compared with Country i, the banks in Country j may continue global-banking activities but engage in foreign sourcing by importing physical capital from abroad. If Country j's endowment of physical capital lies below a sufficiently small threshold, then its banks must obtain sufficient physical capital from abroad to make entrepreneurs indifferent between home and foreign capital and banks, so that they undertake only foreign sourcing. Below an even smaller capital-labor threshold for Country j, maintaining equilibrium requires Country i's to engage in international banking and transmit physical capital to Country j's entrepreneurs.

In Niepmann's model, therefore, differences in the nations' relative endowments of physical capital and labor determine the direction of the net flow of capital between the two countries and hence the direction of international banking or

foreign sourcing (that is, which nation's banks on net lend capital abroad). Differences in national banking efficiencies determine the nature of global banking (that is, which nation's banks on net obtain deposits of capital from abroad).

Kerl and Niepmann (2015) provide an interesting extension of the preceding approach to study banks' choices regarding whether to engage in cross-border trade in loans, in foreign direct investment via acquisitions of foreign firms, and/or in FDI via greenfield entry into host nations. They utilize a model in which banks can lend and borrow in a global interbank market, lend to non-banks domestically and potentially abroad, monitor all loans to assure nonnegative profits, and incur distinctive transactions costs depending on alternative choices among foreign activities. One key implication of their analysis is that interbank lending and foreign lending to non-banks tend to be substitutes. Hence, artificial barriers that restrain lending to non-banks induce more interbank-lending activity can make domestic loan supplies less stable. Another implication is that banks' usage of internal capital markets induces fewer interactions with host-market banks and more exchanges with non-banks, which implies that banks' establishment of foreign affiliates abroad generally is aimed at promoting trade with non-banks. In addition, Niepmann (2016) adapts the Niepmann (2015) framework to a setting with heterogeneous banks. The analysis predicts that more efficient banks are more likely to operate internationally and to maintain larger foreign operations and that fixed costs are fundamental elements influencing the scale of banks' cross-border activities.

An Intra-industry Model of International Trade in Banking Services

An alternative model of cross-border trade in banking services is developed in VanHoose (2013), and the following exposition of this theory follows closely upon this work. The theoretical framework builds directly on the analysis of intra-industry trade exposited by Helpman (1981).

In contrast to Niepmann's framework, crucial assumptions utilized in this international-trade-based model are that banks utilize identical technologies and thereby face the same cost functions and that domestic and foreign banks compete on an equal footing in integrated markets for loans and deposits. Other fundamental differences are that both physical capital and labor are variable factors of production, total desired borrowing and deposits of banks' clienteles are determined independently (exogenously in the model) from outcomes in markets for physical capital and labor, and bank retail markets are imperfectly competitive. As in Niepmann's model, nations' endowments of physical capital and labor can differ.

An element central to interpreting this alternative theory's empirical implications, however, is that like other agents, banks utilize physical capital and labor in the production of retail banking services. Differences in relative factor intensities of

bank's production of loan versus deposit services thereby can perform a crucial role in determining net cross-border flows of these financial services.

A Closed-Economy Setting with Differentiated Banking Services

The model utilizes a Salop (1979)-style monopolistic/monopsonistic-competition similar to the one discussed in Chap. 2. In a closed economy, each one among L borrowers desires one unit of inflation-adjusted credit, and each one among D depositors chooses to make available one unit of real funds. These borrowers and depositors are, along with n banks, arranged uniformly around a unit circle. Each bank possesses service attributes differentiating it from its competitors, and these characteristics are distributed uniformly along a circle with a circumference equal to one. Thus, the separation between attributes, measured by the distance between each of the banks along the unit circle, is equal $1/n$. Borrowers and depositors experience a service-differentiation cost of t per unit of product-characteristic "distance" from the most preferred provider of banking services, and an increase in the number of banks reduces the degree of service differentiation among banks. Consistent with the discussion in Chap. 2, in this setting the profits of bank i are given by

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

where S^i denotes, if a positive (negative) quantity, the bank's net asset (liability) position in a wholesale funding instrument—for instance, interbank loans extended (obtained) at the London interbank offered rate—traded in an interbank lending market at the real rate of return r_S. The function $C_L(L^i, \omega, \delta)$ is the resource cost function applicable to the bank's operations as a lender, where $\omega \equiv$ real wage rate paid to labor, again denoted N. In contrast to the implicit resource-cost functions contemplated elsewhere in this book, this cost function is associated with the loan-services production function $L^i \equiv F_L(N_L, K_L)$, where $N_L \equiv$ lending-specific labor and $K_L \equiv$ lending-specific physical capital. In contrast to Niepmann's model, physical capital is assumed to be a variable input at all banks and other firms in the economy. Likewise, the resource cost function $C_D(D^i, \omega, \delta)$ faced by banks is associated with the deposit-services production function $D^i \equiv F_D(N_D, K_D)$, $\delta \equiv$ real rate of return on physical capital, $N_D \equiv$ deposit-specific labor, and $K_D \equiv$ deposit-specific physical capital. A final assumption is that $C_S(S^i, \omega, \delta) = 0$, which as noted in Chap. 2 is sufficient to ensure that portfolio separation holds true. Thus each bank can borrow or lend in the interbank market at the same marginal rate, resulting in independence of asset and liability choices.

Profit maximization with respect to the balance sheet constraint, $L^i + S^i = D^i$ yields the loan and deposit rates, $r_L = r_S + MC_L(L/n, \omega, \delta) + \frac{t}{n}$, and $r_D = r_S - MC_D(D/n, \omega, \delta) - \frac{t}{n}$. Furthermore, the equilibrium conditions for determination of the market-clearing real wage rate earned by bank-specific labor and real rate of return on physical bank-specific physical capital are given by

$$n[N_L(L/n, \omega, \delta) + N_D(D/n, \omega, \delta)] = N, \text{ and}$$
$$n\left[K_L(L/n, \omega, \delta) + K_D\left(D/n, \omega, \delta\right]\right. = K,$$

where $N_L(L/n, \omega, \delta) = \frac{\partial C_L(L/n, \omega, \delta)}{\partial \omega}$, $N_D(L/n, \omega, \delta) = \frac{\partial C_D(L/n, \omega, \delta)}{\partial \omega}$, $K_L(L/n, \omega, \delta) = \frac{\partial C_L(L/n, \omega, \delta)}{\partial \delta}$, and $K_D(L/n, \omega, \delta) = \frac{\partial C_D(L/n, \omega, \delta)}{\partial \delta}$ are the demands for lending-specific labor, deposit-specific labor, lending-specific physical capital, and deposit-specific physical capital, each of which is identical across banks. Hence, summing each bank's total labor demand and total demand for physical capital across the n banks yields the market input demands, which in equilibrium must match the aggregate endowments of labor, L, and physical capital, K.

In a closed economy, the above relationships determine ω, δ, r_L, and r_D in terms of the number of banks n. In principle, the value of n could be endogenized by imposing a zero-profit condition on the banking system, but this additional long-run condition is not considered here.

A Two-Economy Environment with Traded Banking Services

Consider the following sketch of VanHoose (2013), which extends the above model to a setting with two nations with similar numbers of borrowers and depositors who can utilize the services of either domestic or foreign banks, so that real quantities of loans and deposits move freely between the two countries. Labor and physical capital are immobile across national borders, and fixed numbers of foreign borrowers, foreign depositors, and foreign banks are uniformly distributed in attribute space. Foreign banks face the same loan- and deposit-service production functions and, hence, loan and deposit resource cost functions as those of domestic banks. Foreign borrowers, depositors, and banks join the uniform distributions of domestic borrowers, depositors, and banks, and intra-industry international trade thereby characterizes both sets of retail banking markets.

Although the advent of international trade in banking services leaves loan and deposit markets imperfectly competitive, suppose domestic and foreign banks compete head to head to provide lending and deposit services. The international interbank market likewise is integrated, and the resulting real interest parity yields the market-clearing real wholesale interbank rate r_S. Proceeding through domestic and foreign banks' optimization problems along lines analogous to the closed-economy case will yield integrated-markets, no-arbitrage solutions for the market loan and deposit rates. In equilibrium, the sum of global borrowing must equal the

aggregate quantity of loans provided by all domestic and foreign banks. Likewise, total deposits across the two nations must be equivalent to the aggregate amount of deposits issued across all domestic and foreign banks. Taken together with equilibrium conditions for determination of the market-clearing real wage rates earned by immobile labor and real rates of return on immobile physical capital in the two countries, the two no-arbitrage conditions for bank loan and deposit markets, the two summing-up conditions yield *simultaneous* solutions for domestic and foreign lending and deposits and domestic and foreign wage rates and rates of return on physical capital. Loan and deposit rates adjust to market-clearing levels.

A standard Heckscher-Ohlin-style outcome emerges from this model of bank intra-industry trade: Relative intensities of labor and physical capital in the production of loan and deposit services and relative endowments of these factors determine flows of trade in these banking services. Given the relatively higher endowment of banking-applicable physical capital in the domestic nation assumed in the situation graphed in the figures, the domestic real rate of return on physical capital utilized in banking is lower and the domestic real wage rate earned by bank-employed labor is higher than in the foreign country. On the one hand, domestic banks respond by exporting deposit services, so that a portion of foreign deposit funds flows to the domestic nation. On the other hand, foreign banks respond by exporting lending services, so that a portion of total credit flowing to domestic borrowers is obtained from foreign banks.

Consequently, the global wholesale interbank funds market is a crucial mechanism for enabling the flows of funds between the two nations consistent with attainment of the profit-maximizing banking-service trade flows. In this instance, foreign banks compensate for the net outflow of deposit funds from the foreign country, which otherwise would render infeasible the profit-maximizing net exporting of lending services, by borrowing funds from domestic banks in the interbank market. A fundamental implication of the model, therefore, is that a smoothly functioning global interbank funds market is a necessary condition for fulfillment of profit-maximizing intra-industry bank trade flows. Hence, to the extent that nations view broad-based international trade in banking services as socially desirable—see, for instance, Groutzinis (2005)—a strong foundation for liquid operations of global interbank markets is a key prerequisite.

The model further indicates that if providing deposit services is physical-capital-intensive while lending is labor-intensive and one nation—say, the domestic nation—has a relatively greater endowment of physical capital while the other—say, the foreign nation—has a relatively larger endowment of labor, then the domestic nation will export deposit services while the foreign nation exports lending services. Hence, the domestic nation will experience inflows of both deposit funds and credit, while the foreign nation will experience outflows of both. Furthermore, the domestic country's banking system will be a net seller of wholesale funds in the global interbank market, while the foreign country's banking system will be a net borrower of wholesale funds.

More generally, however, the theory's fundamental implication is that relative intensities of utilization of factors of production by nations' banking systems

should, together with relative factor endowments, be key determinants of net cross-border flows of banking services. Thus, differences in relative factor intensities rather than—or perhaps alongside—intermediation efficiencies influence international banking flows.

Evidence on the Empirical Relevance of Trade-Based Theories of International Banking

The idea that traditional resource-based economic models might help to explain cross-border banking flows has only recently been explored theoretically, but a few handful of researchers already have thought to confront basic predictions of the international-trade literature with banking data. For instance, in a study of bilateral flows of banking services among 141 nations, Wengel (1995) finds evidence that the intra-industry trade incentives best explain the observed flows. Likewise, in the context of analysis of OECD data on financial services trade on the part of 29 nations, Webster and Hardwick (2005) conclude that there is evidence of "significant volumes" of intra-industry trade in financial services.

The explicit resource-based models of cross-border banking offered by Niepmann (2015) and VanHoose (2013) provide more specificity regarding relevant explanatory variables and predicted effects on l banking flows. In the case of Niepmann's theory, the explanatory variables of interest are national (human-capital-adjusted) ratios of aggregate physical capital resources to quantities of labor and measures of bank efficiency. Niepmann's proxy for the latter variable is each nation's ratio of total bank overhead expense (reported in the World Bank's Financial Structure Database) to total bank assets. For the 25 nations included in her empirical analysis, Niepmann utilizes data on these variables in 2000. To prevent endogeneities between these variables and her dependent variables, she relates these two explanatory variables to 2005 observations of the dependent variables.

The dependent variables observed in 2005 that Niepmann considers are banks' consolidated cross-border (including foreign affiliates) positions available in the Consolidated Banking Statistics dataset maintained by the Bank for International Settlements and data on cross-border positions of German banks in the Auslandsstatus-Report provided by the Deutsche Bundesbank. As additional explanatory variables Niepmann includes contemporaneous differences in returns on loans, distance and language variables to proxy frictions, measures of financial-sector openness, and other controls GDP, GDP per capita, a measure of the degree of protection of property rights, and a dummy variable taking into account banking crises.

Support for essential aspects of Niepmann's theory is provided by the fact that, consistent with her model's predictions, national ratios of physical capital to labor have (with varying degrees of statistical significance) generally positive effects on bank foreign asset positions and negative effects on foreign liability positions. Measured effects of the overhead-to-assets efficiency proxy, which tend to be

more strongly related to the dependent variables, indicate that more assets are held abroad by residents of nations with more efficient banking systems and that nations with more efficient banking systems raise more funds in countries with less efficient banking systems. The efficiency measure, Niepmann finds, has a stronger effect on banks' foreign liabilities than on their foreign assets.

The alternative theory discussed above also has testable empirical implications. Dia and VanHoose (2017) explore the most essential, which is that varying degrees of factor intensities should help to explain observed loan and deposit flows across national borders. They focus on physical capital intensities, which they estimate by dividing the total value of nations' banks' fixed and intangible assets by the number of bank employees for 27 nations over the interval between 2001 and 2012. Their overall measure of physical capital intensities for each nation's banking system in a given year is constructed as a weighted average using banks' shares of total national deposits as weights, although they find very similar results using banks' shares of total loans as weights.

The intra-industry-trade-based theory indicates that net exports of retail loans and deposits of different countries available (after removing interbank transactions) from the cross-border banking data provided by the Bank for International Settlements, should vary systematically with those nations' physical capital intensities. Naturally, for a number of countries, net exports of such banking services are negative, so Dia and VanHoose's empirical specifications are expressed in terms of year-to-year first differences in these variables. Hence, their econometric models actually relate annual *changes* in nations' banks' net loan and deposit exports to *changes* in intensities of the banks' utilization of physical capital as well as to other variables that might affect international banking flows, such as levels of GDP, variations of nations' open-market interest rates from the U.S. rate, and the magnitudes of countries' external balance positions. After taking into account the separate effects of these additional explanatory variables, Dia and VanHoose find evidence of economically and statistically significant negative relationships between the changes in the capital-intensity measures and changes in net exports of both loans and deposits, with the results indicative of a relatively greater intensity of utilization of physical capital for deposits than for loans. Separate empirical estimations relating changes in the levels of loan and deposit exports and imports—rather than the net of the two quantities—to changes in physical capital intensities verify these conclusions. Consistent with theory and the results for net exports, levels of exports and imports of retail bank loans and deposits respond in opposite directions to changes in capital intensities.

References

Adams-Kane, Jonathan, Yuequing Jia, and Jamus Jerome Lim. 2015. Global transmission channels for international bank lending in the 2007–2009 financial crisis. *Journal of International Money and Finance* 56: 97–113.

Aliber, Robert. 1984. International banking: A survey. *Journal of Money, Credit, and Banking* 16: 661–678.

Battilossi, Stefano. 2006. The determinants of multinational banking during the first globalization, 1870–1914. *European Review of Economic History* 10: 361–388.

Beccalli, Elena, and Pascal Frantz. 2009. M&A operations and performance in banking. *Journal of Financial Services Research* 36: 203–226.

Beck, Thorsten. 2002. Financial development and international trade: Is there a link? *Journal of International Economics* 57: 107–131.

Beck, Thorsten, Aslie Demirgüç-Kunt, and Maria Soledad Martinez Pería. 2011. Bank financing for SMEs: Evidence across countries and bank ownership types. *Journal of Financial Services Research* 39: 35–54.

Berger, Allen, Robert DeYoung, Hesna Genay, and Gregory Udell. 2000. Globalization of financial institutions: Evidence from cross-border banking performance. *Brookings-Wharton Papers on Financial Services* 3: 23–120.

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.

Bremus, Franziska. 2015. Cross-border banking, bank market structures, and market power: Theory and cross-country evidence. *Journal of Banking and Finance* 50: 242–259.

Bremus, Franziska, and Marcel Fratzscher. 2014. Drivers of structural change in cross-border banking since the global financial crisis. DIW Berlin Discussion Paper 1411, September.

Buch, Claudia. 2003. Information or regulation? What drives the international activities of commercial banks? *Journal of Money, Credit, and Banking* 35: 851–869.

Buch, Claudia, Cathérine Koch, and Michael Koetter. 2011. Size, productivity, and international banking. *Journal of International Economics* 85: 329–334.

Cetorelli, Nicola, and Linda Goldberg. 2016. Organizational complexity and balance sheet management in global banks. National Bureau of Economic Research Working Paper No. 22169, April.

Claessens, Stijn, and Neeltje van Horen. 2012. Being a foreigner among domestic banks: Asset or liability? *Journal of Banking and Finance* 36: 1276–1290.

Claessens, Stijn, and Neeltje van Horen. 2014a. Location decisions of foreign banks and competitor remoteness. *Journal of Money, Credit, and Banking* 46: 145–170.

Claessens, Stijn, and Neeltje van Horen. 2014b. Foreign banks: Trends and impact. *Journal of Money, Credit, and Banking* 46: 295–326.

Claessens, Stijn, and Neeltje van Horen. 2015. The impact of the global financial crisis on banking globalization. *IMF Economic Review* 63: 869–918.

Clare, Andrew, Mohamed Azzim Gulamhussen, and Carlos Pinheiro. 2013. What factors cause foreign banks to stay in London? *Journal of International Money and Finance* 32: 739–761.

Clerides, Sofronis, Manthos Delis, and Sotirios Kokas. 2015. A new data set on competition in national banking markets. *Financial Markets, Institutions, and Instruments* 24: 267–311.

Correa, Ricardo. 2009. Cross-border bank acquisitions: Is there a performance effect? *Journal of Financial Services Research* 36: 169–197.

Curry, Elisa, Justin Fung, and Ian Harper. 2003. Multinational banking: Historical, empirical, and case perspectives. In *Handbook of International Banking*, ed. Andrew Mullineux and Victor Murinde, 27–59. Cheltenham, UK: Edward Elgar.

De Blas, Beatriz, and Katheryn Russ. 2010. FDI in the banking sector. Unpublished Manuscript, Universidad Autónoma de Madrid and University of California, Davis, and NBER.

De Haas, Ralph, and Iman van Lelyveld. 2010. Internal capital markets and lending by multinational bank subsidiaries. *Journal of Financial Intermediation* 19: 1–25.

Degryse, Hans, Olena Havrylchyk, Emilia Jurzyk, and Sylwester Kozak. 2012. Foreign bank entry, credit allocation, and lending rates in emerging markets: Evidence from Poland. *Journal of Banking and Finance* 36: 2949–2959.

Dia, Enzo, and Lorenzo Menna. 2016. Productivity shocks, capital intensities, and bank interest rates. *Journal of Macroeconomics* 48: 155–171.

Dia, Enzo, and David VanHoose. 2017. Capital intensities and international trade in banking services. *Journal of International Financial Markets, Institutions, and Money* 46: 54–69.
Dietrich, Diemo, and Uwe Vollmer. 2010. International banking and liquidity allocation: Cross-border financial services versus multinational banking. *Journal of Financial Services Research* 37: 45–69.
Do, Quy-Toan, and Andrei Levchenko. 2007. Comparative advantage, demand for external finance, and financial development. *Journal of Financial Economics* 86: 796–834.
García-Herrero, Alicia, and Francisco Vázquez. 2013. International diversification gains and home bias in banking. *Journal of Banking and Finance* 37: 2560–2571.
Goldberg, Linda. 2009. Understanding banking sector globalization. *IMF Staff Papers* 56: 171–197.
Gormley, Todd. 2010. The impact of foreign bank entry in emerging markets: Evidence from India. *Journal of Financial Intermediation* 19: 26–51.
Groutzinis, Apostolos. 2005. International trade in banking services and the role of the WTO: Discussing the legal framework and policy objectives of the General Agreement on Trade in Services and the current state of play in the Doha Round of trade negotiations. *International Lawyer* 38: 877–914.
Havrylchyk, Olena. 2012. The effect of foreign bank presence on firm entry and exit in transition economies. *Journal of Banking and Finance* 36: 1710–1721.
Helpman, Elhanan. 1981. International trade in the presence of product differentiation, economies of scale, and monopolistic competition: A Chamberlin-Heckscher-Ohlin approach. *Journal of International Economics* 11: 305–340.
Huertas, Thomas. 1990. U.S. multinational banking: History and prospects. In *Banks as Multinationals*, ed. Geoffrey Jones, 248–267. London: Routledge.
Jaud, Melise, Madina Kukenova, and Martin Strieborn. 2014. Financial development and sustainable exports: Evidence from firm-product data. *The World Economy* 38: 1090–1114.
Jeon, Bang Nam, Maria Pía Olivero and Ji Wu. 2011. Do foreign banks increase competition? Evidence from emerging Asian and Latin American banking markets. *Journal of Banking and Finance* 35: 856–875.
Kerl, Cornelia, and Friederike Niepmann. 2015. What determines the composition of international bank flows? *IMF Economic Review* 63: 792–829.
Kleimeier, Stefanie, Harald Sander, and Sylvia Heuchemer. 2013. Financial crises and cross-border banking: New evidence. *Journal of International Money and Finance* 32: 884–913.
Kodongo, Odongo, Dinah Natto, and Nicholas Biekpe. 2015. Explaining cross-border bank expansion in East Africa. *Journal of International Financial Markets, Institutions, and Money* 36: 71–84.
Lehner, Maria. 2009. Entry mode choice of multinational banks. *Journal of Banking and Finance* 33: 1781–1792.
Lewis, Mervyn, and Kevin Davis. 1987. *Domestic and International Banking*. Oxford, UK: Philip Allan.
Li, Donghui, Fariborz Moshirian, and Ah-Boon Sim. 2003. The determinants of intra-industry trade in insurance services. *Journal of Risk and Insurance* 70: 269–287.
Manlagñit, Maria Chelo. 2011. The economic effects of foreign bank presence: Evidence from the Philippines. *Journal of International Money and Finance* 30: 1180–1194.
Minoiu, Camelia, and Javier Reyes. 2013. A network analysis of global banking. *Journal of Financial Stability* 9: 168–184.
Moshirian, Fariborz. 1994. Trade in financial services. *World Economy* 17: 347–364.
Moshirian, Fariborz, Donghui Li, and Ah-Boon Sim. 2005. Intra-industry trade in financial services. *Journal of International Money and Finance* 24: 1090–1107.
Müller, Oliver, and André Uhde. 2013. Cross-border lending: Empirical evidence on new determinants from OECD markets. *Journal of International Financial Markets, Institutions, and Money* 23: 136–162.

Mulyaningsih, Tri, Anne Daly, and Royana Miranti. 2015. Foreign participation and banking competition: Evidence from the Indonesian banking industry. *Journal of Financial Stability* 19: 70–82.

Neu, C.R. 1988. International trade in banking services, Chapter 10. In *Issues in US-EC Trade Relations*, ed. Robert Baldwin, Carl Hamilton, and Andre Sapir, 245–269. Chicago, IL: University of Chicago Press.

Niepmann, Friederike. 2015. Banking across borders. *Journal of International Economics* 96: 244–265.

Niepmann, Friederike. 2016. Banking across borders with heterogeneous banks. International Finance Discussion Paper No. 1177, Board of Governors of the Federal Reserve System, August.

Popov, Alexander, and Steven Ongena. 2011. Interbank market integration, loan rates, and firm leverage. *Journal of Banking and Finance* 35: 544–559.

Reinhardt, Dennis, and Steven Riddiough. 2015. The two faces of cross-border banking flows. *IMF Economic Review* 63: 751–791.

Rughoo, Aarti, and Nicholas Sarantis. 2014. The global financial crisis and integration in European retail banking. *Journal of Banking and Finance* 40: 28–41.

Sagari, Silvio. 1989. International trade in financial services. World Bank Working Paper No. 134, January.

Salop, Steven. 1979. Monopolistic competition with outside goods. *Bell Journal of Economics* 10: 141–156.

Shirota, Toyoichiro. 2015. What is the major determinant of cross-border banking flows? *Journal of International Money and Finance* 53: 137–147.

Sturm, Jan-Egbert, and Barry Williams. 2010. What determins differences in foreign bank efficiency? *Journal of International Financial Markets, Institutions, and Money* 20: 284–309.

Temesvary, Judit. 2014. The determinants of U.S. banks' international activities. *Journal of Banking and Finance* 44: 233–247.

Temesvary, Judit. 2015. Foreign activities of U.S. banks since 1997: The roles of regulations and market conditions in crises and normal times. *Journal of International Money and Finance* 56: 202–222.

VanHoose, David. 2013. A model of international trade in banking services. *Open Economies Review* 24: 613–625.

Webster, Allan, and Philip Hardwick. 2005. International trade in financial services. *The Service Industries Journal* 25: 721–746.

Wengel, Jan. 1995. International trade in banking services. *Journal of International Money and Finance* 14: 47–64.

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.

Banks 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

© Springer-Verlag GmbH Germany 2017

D. VanHoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-662-54326-9_5

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 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. 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 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 mergers imply that acquired banks should share features that identify them as 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. 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% 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. They suggest 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 (2009) 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.

Egger and Hahn (2010) offer contrasting conclusions. They seek to control for treatment effects in evaluating bank mergers in Austria between 1996 and 2002 via a propensity-score matching approach. Their results indicate significant cost-efficiency gains from mergers, particularly for smaller banks, that contribute to better profit performance, ceteris paribus.

Although Cheng et al. (1989) also examine the characteristics of acquisition targets, which they find tend to perform below 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 (2009) 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 8000 banking institutions, of which more than 1400 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 study of more than 150 EU 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 EU banking mergers between 1997 and 2002.

Beccalli and Frantz (2013) analyze 777 mergers and acquisitions involving EU banks and 312 acquired institutions. They also provide evidence that acquired institutions tend to be more highly levered, as well as less liquid and less efficient. They suggest that this finding is consistent acquiring banks having more skillful managers who can improve the capital and liquidity positions and cost efficiencies of acquired institutions.

One potential pitfall in assessing profitability rationales for mergers in recent years is that many mergers have been encouraged by regulators seeking to subsume weaker institutions within stronger banks. Bauer et al. (2009), for instance, find in a study of U.S. credit union mergers promoted by regulators between 1994 and 2004 that in such cases the main beneficiaries of the mergers typically are stakeholders of the acquired institutions and the deposit insurance system.

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 5000 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 1000 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. In addition, Caizza et al. (2012) examine 1156 domestic and 328 cross-border bank merger deals between 1988 and 2006, and find that acquired firms tend to be less cost-efficient. They also conclude, however, that acquirers seek institutions displaying higher risk profiles, particularly

in the case of banks acquired abroad, a conclusion that potentially reinforces the view that diversification of risks influences mergers.

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 $(1-q)\left(r_S - C_S^1\right) - C_D^1$, 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.

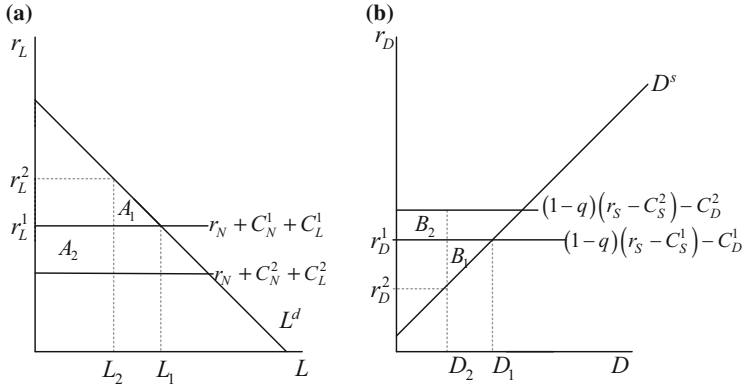

Fig. 5.1 Effects of industry consolidation in initially perfectly competitive loan and deposit markets

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 Chap. 3, 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, and C_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.

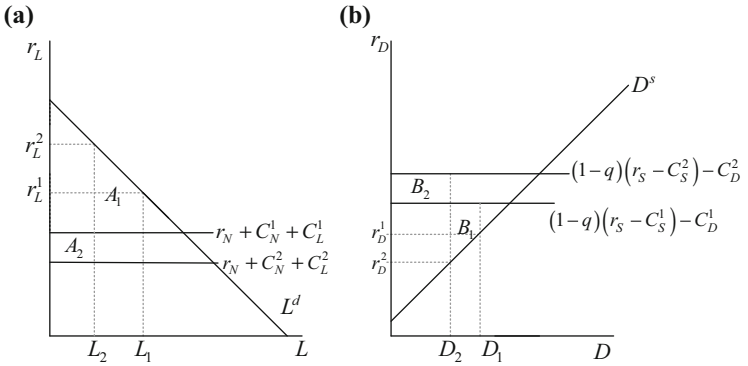

Fig. 5.2 Effects of industry consolidation in initially imperfectly competitive loan and deposit markets

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 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 consolidation arises if it 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.

Banking industry consolidation again generates 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 theoretical net market impacts of bank mergers are ambiguous. Thus, assessing these net impacts is an empirical issue.

Mergers and Market Power

Does banking consolidation contribute to increased market power? As discussed later in this chapter, regulatory merger guidelines presume that mergers boost market concentration and, ultimately, market power. Evidence provided by Adams et al. (2009), however, indicates that banking markets in which mergers yield immediately higher concentration levels eventually tend to experience lower concentration within the following three to 5 years. Their analysis finds less clear-cut longer-term effects on concentration in markets that experience larger immediate *changes* in concentration following mergers.

Most research seeking to investigate the relationship between bank mergers and market power 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% 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. Based upon examination of data from more than 3000 business borrowers from banks participating in large mergers, Fraser et al. (2011) conclude that borrowers' stock prices are adversely affected, which they interpret as evidence of increased market power for consolidated banking firms created by the mergers. Hankir et al. (2011) reach similar conclusions about market-power effects of mergers based upon analysis of stock returns of acquirers, targets, and peers affected by 600 North American and European bank mergers between 1990 and 2008.

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. They also find that the most pronounced effect of mergers is on 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 examinations of monthly U.S. banking data encompassing mergers between 1997 and 2006, Craig and Dinger (2009) find evidence of downward pressure on rates paid on transactions deposits, and Dinger (2015) concludes that the frequency of deposit rate adjustments also declines. 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.

Considerable evidence indicates that impacts of banking consolidation on borrowers and depositors differ depends on whether they are customers of an acquiring bank or an acquired bank. Carow et al. (2006) analyze data on more than 2000 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 1500 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 through 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%. 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% 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) and 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 7000 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-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 (2009) 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. Furthermore, in a study of more than 700 bank merger deals by European Union acquirers between the early 1990s and mid-2000s, Beccalli and Frantz (2009) find that

improved cost efficiencies generated by the mergers for consolidated institutions were accompanied by deteriorations in profit performance, so that the main merger beneficiaries were bank customers.

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 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 examine how measured productive efficiency of non-merging banks responded to consolidations of other banks. 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% 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 antitrust analysis. In practice, regulators and the Department of Justice have chosen to utilize 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 market, the Federal Reserve uses the Herfindahl-Hirschman index (HHI), the sum of squared percentage market shares of each of all banks. It computes both pre-merger and post-merger HHIs and 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 $\sum_{i=1}^{n} (S_i)^2$, where there are n banks in the relevant market and S_i is the market share of bank i. After the merger of two banks 1 and 2, it is $\sum_{i=1}^{n} (S_i)^2 - (S_1)^2 - (S_2)^2 + (S_1 + S_2)^2 = \sum_{i=1}^{n} (S_i)^2 + 2S_1S_2$; 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 influence an assessment of a proposed merger.

Merger Screening

In evaluating planned consolidations of nonfinancial firms, the U.S. Department of Justice has since 2010 considered a market to be unconcentrated if its HHI is less than 1500, to be moderately concentrated if its HHI is between 1500 and 2500, and to be highly concentrated if its HHI exceeds 2500. Figure 5.3 suggests that relying on these classifications of market concentration would imply that many U.S. banking markets, as defined by metropolitan statistical areas, at least, are moderately concentrated. If one had depended on pre-2010 Justice Department guidelines, however, in which the critical HHI value for a highly concentrated market was 1800, only since the late 2000s would this judgment have been rendered.

As the figure indicates, the average HHI value for U.S. MSAs trended downward between the early 1990s and mid-2000s. During more recent years, however, the MSA has jumped upward.

A number of 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 2500

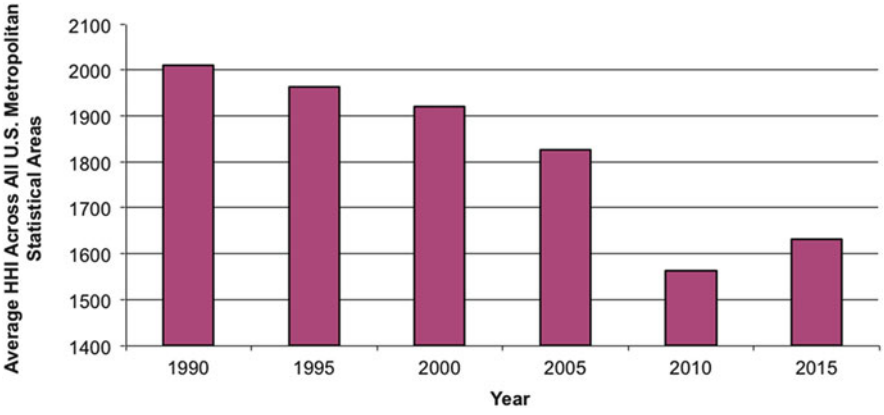

Fig. 5.3 Average Herfindahl-Hirshman Index values in U.S. metropolitan statistical areas [Sources: Pilloff (2009) and Federal Reserve Bank of St. Louis (2017)]

and the change in the HHI resulting from the merger exceeds 100. 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% weight to the deposits of commercial banks but placing only a 50% 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 2500. 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 a second screen, called Screen B, applies. It uses the 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, the Federal Reserve's 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 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 use to assess the prospects for approval of a 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, the finding of a positive long-run cross-price 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% above within a 1-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 official 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% 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 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 the more recent years.

Nevertheless, Gilbert and Zeretsky 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 6-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 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 Chap. 3, 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) evaluates 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 deposit rates but concludes that market loan rates are better explained by the number of competitors than by the HHI. Carbó-Valverde et al. (2009) consider a cross section of banking data from more than 1900 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. Bolt and Humphrey (2015) evaluate whether competition measures including the HHI, Lerner Index, and H-Statistic are consistent with a constructed revenue-cost competition frontier for more than 2600 banks participating in the U.S. consumer loan market between 2008 and 2010 and conclude that the HHI is a particularly weak measure of the intensity of competition.

Azar et al. (2015) argue that using standard HHI measures that focus on concentration among banking firms potentially are less useful than alternative measures that capture concentration of bank ownership. They provide evidence that banks' deposit account maintenance fees and deposit rate spreads are unrelated to standard HHI measures. The authors suggest that in contrast, these fees and spreads do depend in SCP-predicted ways on a "generalized HHI," that takes into account both common ownership of banks by the same investors and cross-ownership of banking institutions.

Implications of Endogenous Sunk Fixed Costs

As discussed in Chap. 3, 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 1800 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?

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 contributions that have progressed in extending several advances in industrial organization to the realm of banking markets. One area that 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 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% 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? 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 merger announcements 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 2006). 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. Demand estimations utilized for simulating merger outcomes can then become 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 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% rise in the loan-rate margin) and a 10-basis-point decrease in the market loan rate (which implies a 37% increase in the deposit-rate margin). Molnár notes

that these estimated impacts are relatively small in light of the significant market concentration—about 70%—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 in McIntosh (2002). This author 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.

A number of event studies of bank mergers exist. 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.

References

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 Govenors of the Federal Reserve System.

Adams, Robert, Richard Johnson, and Steven Pilloff. 2009. Market structure after horizontal mergers: Evidence from the banking industry. *Review of Industrial Organization* 35: 217–231.

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.

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.

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 Bar Association. 2007. Bank Mergers and Acquisitions Handbook. Section on Antitrust Law.

Ashton, John, and Khac Pham. 2007. Efficiency and price effects of horizontal bank mergers. Universtiy 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.

Azar, José, Sahil Raina, and Martin Schmalz. 2015. Ultimate ownership and bank competition. Unpublished manuscript, Charles River Associates and University of Michigan.

Bauer, Keldon, Linda Miles, and Takeshi Nishikawa. 2009. The effect of mergers on credit union performance. *Journal of Banking and Finance* 33: 2267–2274.

Beccalli, Elena, and Pascal Frantz. 2009. M&A operations and performance in banking. *Journal of Financial Services Research* 36: 203–226.

Beccalli, Elena, and Pascal Frantz. 2013. The determinants of mergers and acquisitions in banking. *Journal of Financial Services Research* 43: 265–291.

Becher, David, and I.I. Terry Campbell. 2005. Interstate banking deregulation and the changing nature of bank mergers. *Journal of Financial Research* 28: 1–20.

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, Sigbjorn 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.

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 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 Astrid Dick. 2007. Entry into banking markets and early-mover advantage. *Journal of Money, Credit, and Banking* 39: 775–807.

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, 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, 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.

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.

Bolt, Wilko, and David Humphrey. 2015. Assessing bank competition for consumer loans. *Journal of Banking and Finance* 61: 127–141.

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.

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.

Caizza, Stefana, Andrew Clare, and Alberto Franco Pozzolo. 2012. What do bank acquirers want? Evidence from worldwide bank M&A targets. *Journal of Banking and Finance* 36: 2641–2659.

Campa, José Manuel, and Ignacio Hernando. 2006. M&As performance in the European financial industry. *Journal of Banking and Finance* 30: 3367–3392.

Carbó-Valverde, Santiago, David Humphrey, Joaquin Maudos, and Philip Molyneux. 2009b. Cross-country comparisons of competition and pricing power in European banking. *Journal of International Money and Finance* 28: 115–134.

Carow, Kenneth, Edward Kane, and Rajesh Narayanan. 2006. How have borrowers fared in banking megamergers? *Journal of Money, Credit, and Banking* 38: 821–836.
Cheng, David, Benton Gup, and Larry Wall. 1989. Financial determinants of bank takeovers. *Journal of Money, Credit, and Banking* 21: 524–536.
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.
Craig, Ben, and Valeriya Dinger. 2009. Bank mergers and the dynamics of deposit interest rates. *Journal of Financial Services Research* 36: 111–133.
Cybo-Ottone, Alberto, and Maurizio Murgia. 2000. Mergers and shareholder wealth in European banking. *Journal of Banking and Finance* 24: 831–859.
DeLong, Gayle. 2001. Stockholder gains from focusing versus diversifying bank mergers. *Journal of Financial Economics* 59: 221–252.
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.
Dick, Astrid. 2007. Market size, service quality, and competition in banking. *Journal of Money, Credit, and Banking* 39: 49–81.
Dick, Astrid. 2008. Demand estimation and consumer welfare in the banking industry. *Journal of Banking and Finance* 32: 1661–1676.
Dinger, Valeriya. 2015. Bank mergers and deposit rate rigidity. *Journal of Financial Services Research* 47: 27–56.
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.
Egger, Peter, and Franz Hahn. 2010. Endogenous bank mergers and their impact on banking performance: Some evidence from Austria. *International Journal of Industrial Organization* 28: 155–166.
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* 25: 259–281.
Erel, Isil. 2009. The effect of bank mergers on loan prices: Evidence from the United States. *Review of Financial Studies* 29: 413–448.
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: 897–928.
Federal Reserve Bank of St. Louis, 2017, Cassidi Database (https://cassidi.stlouisfed.org/).
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.
Fraser, Donald, James Kolari, Seppo Pynnönen, and Kyle Tippens. 2011. Market power, bank megamergers, and the welfare of bank borrowers. *Journal of Financial Research* 34: 641–658.
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.
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.
Hadlock, Charles, Joel Houston, and Michael Ryngaert. 1999. The role of managerial incentives in bank acquisitions. *Journal of Banking and Finance* 23: 221–249.
Hall, Robert. 2007. Book review: Lectures on antitrust economics. *Journal of Economic Literature* 45: 1066–1070.
Hankir, Yassin, Christian Rauch, and Marc Umber. 2011. Bank M&A: A market power story? *Journal of Banking and Finance* 35: 2341–2354

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, and Steven Pilloff. 2009. Acquisition targets and motives in the banking industry. *Journal of Money, Credit, and Banking* 41: 1167–1187.

Hannan, Timothy, and Stephen Rhoades. 1987. Acquisition targets and motives: The case of the banking industry. *Review of Economics and Statistics* 69: 67–74.

Hauswald, Robert, and Robert Marquez. 2006. Competition and strategic information acquisition in credit markets. *Review of Financial Studies* 19: 967–1000.

Heitfield, Erik, and Robin Prager. 2004. The geographic scope of retail deposit markets. *Journal of Financial Intermediation* 25: 37–55.

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.

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.

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.

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.

Jackson, William. 1992. Is the market well defined in bank merger and acquisition analysis? *Review of Economics and Statistics* 74: 655–661.

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.

Karceski, Jason, Steven Ongena, and David Smith. 2005. The impact of bank consolidation on commercial borrower welfare. *Journal of Finance* 60: 2043–2082.

Kim, Moshe, Doron Kliger, and Bent Vale. 2003. Estimating switching costs: The case of banking. *Journal of Financial Intermediation* 12: 25–56.

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.

Koetter, J.W.B., F. Bos, J.W. Heid, C.J.M. Kool Kolari, and D. Porath. 2007. Accounting for distress in bank mergers. *Journal of Banking and Finance* 31: 3200–3217.

Kwast, Myron, Martha Starr-McCluer, and John Wolken. 1997. Market definition and the analysis of antitrust in banking. *Antitrust Bulletin* 42: 973–995.

Liang, Nellie, and Stephen Rhoades. 1988. Geographic diversification and risk in banking. *Journal of Economics and Business* 40: 271–284.

McIntosh, James. 2002. A welfare analysis of Canadian chartered bank mergers. *Canadian Journal of Economics* 35: 457–475.

Molnár, József. 2008. Market power and merger simulation in retail banking. Bank of Finland Research Discussion Paper #4-2008.

Palia, Darius. 1993. The managerial, regulatory, and financial determinants of bank merger premiums. *Journal of Industrial Economics* 41: 91–102.

Park, Kwangwoo, and George Pennacchi. 2008. Harming depositors and helping borrowers: The disparate impact of bank consolidation. *Review of Financial Studies* 22: 1–40.

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.

Patti, Di, Emilia Bonaccorsi, and Giogio Gobbi. 2007. Winners or losers? The effects of banking consolidation on corporate borrowers. *Journal of Finance* 62: 669–695.

Perloff, Jeffrey, Larry Karp, and Amos Golan. 2007. *Estimating Market Power and Strategies*. Cambridge, UK: Cambridge University Press.

Peters, Craig. 2006. Evaluating the performance of merger simulation: Evidence from the U.S. airline industry. *Journal of Law and Economics* 49: 627–649.

Pilloff, Steven. 2009. The banking industry. In *The Structure of American Industry*, ed. James Brock, 12th ed. Upper Saddle River, NJ: Pearson.

Radecki, Lawrence. 1998. The expanding geographic reach of retail banking markets. *Federal Reserve Bank of New York Economic Policy Review* 4: 15–34.

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. 1997. Have barriers to entry in retail commercial banking disappeared? *Antitrust Bulletin* 42: 997–1013.

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.

Rivard, Richard, and Christopher Thomas. 1997. The effect of interstate baning on large bank holding company profitability and risk. *Journal of Economics and Business* 49: 61–76.

Sapienza, Paola. 2002. The effects of banking mergers on loan contracts. *Journal of Finance* 57: 329–367.

Scott, Jonathan, and William Dunkelberg. 2003. Bank mergers and small firm financing. *Journal of Money, Credit, and Banking* 35: 999–1017.

Sharpe, Steven. 1990. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *Journal of Finance* 49: 1069–1087.

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.

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.

Stigler, George. 1955. Introduction. In *Business Concentration and Price Policy*, National Bureau of Economic Research. Princeton, NJ: Princeton University Press.

Stigler, George, and Robert Sherwin. 1985. The extent of the market. *Journal of Law and Economics* 29: 555–585.

Walker, Michael. 2005. Merger simulation models: Useful or just plain dangerous. Presentation Notes, Global Competition Law Centre Seminar, College of Europe, October 21.

Walter, Ingo. 2004. *Mergers and Acquisitions in Banking and Finance*. Oxford: Oxford University Press.

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.

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.

Whinston, Michael. 2006. *Lectures on Antitrust Economics*. Cambridge, MA: MIT Press.

Whinston, Michael. 2007. Antitrust policy toward horizontal mergers. In *Handbook of Industrial Organization*, Volume 3, 2371–2440. Amsterdam: Elsevier North Holland.

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. Cole and White (2012) find evidence that proxy measures for components of *CAMELS* ratings provide strong explanatory power for bank failures observed in 2009.

This chapter reviews the rationales that have motivated 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

© Springer-Verlag GmbH Germany 2017

D. VanHoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-662-54326-9_6

constrain the activities of individual banks and impinge on the structure of the banking industry.

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 share risks 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 examine an 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) $u'(c^*_{1,1}) = \rho R u'(c^*_{2,2})$, 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) $pc^*_{1,1} + (1-p)\frac{c^*_{2,2}}{R} = 1$ (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 $\frac{u'(c^*_{1,1})}{u'(c^*_{2,2})}$ 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, 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. If 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 several attractive features. First, it motivates how an intermediary might come into existence—namely, as a mechanism for implementing a risk-sharing contract among agents with differential consumption timing. Second, it provides a rationale for bank runs as 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 additional background assumptions 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 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, 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 via "deposit insurance" can create moral hazard problems that must be addressed by additional, likely non-costless, interventions. After all, the expected value of a 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. 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. One interesting exception is a contribution by Iyer and Puri (2008). 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. They identify two other 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.

Another exception is the study by Goedde-Menke et al. (2014), who conducted surveys of groups of more than 600 bank depositors before, near, and after the peak of the recent global financial crisis. These surveys revealed that depositors' information about insurance of their funds and their views about security of their deposits varied across individuals and over time. Indeed, Goedde-Menke found that the greatest risk of a bank run was not near the crisis peak but instead after that peak.

Additional exceptions are the experimental analyses of Schotter and Yorulmazer (2009) and Kiss et al. (2012). The former authors find in a laboratory setting that when more information is available as a crisis develops, the likelihood of a run decreases and that even limited forms of deposit insurance can reduce the likelihood of a run. The latter researchers conclude that while higher levels of deposit-insurance provision help to reduce the likelihood of bank runs when agents cannot observe others' actions, observability of agents' actions tends to offset this effect.

These authors' 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 heterogeneities across time, such as the Diamond-Dybvig assumption that depositors are endowed with predispositions to withdraw funds in different periods. Otherwise, depositors' characteristics and information sets otherwise are identical in typical Diamond-Dybvig-style models. Another 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 and policymakers.

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. 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.

Hainz et al. (2013) examine the effects of greater bank competition on the choice between screening loans versus requiring collateral to mitigate adverse selection problems. In their theoretical model, costs of screening increase with greater distance between banks and borrowers. An increase in the number of banks naturally reduces screening expenses and thereby boosts the amount of screening while reducing collateralization of loans—an implication that Hainz et al. find receives support in data on more than 4900 loans in 70 nations between 2000 and 2005.

Gomez and Ponce (2014) also offer theoretical analysis of the relationship between bank competition and loan quality in light adverse-selection problems. In their model, similarly to the analysis of Hainz et al. (2013), an increase in the number of competing banks initially implies more screened loans and hence higher loan quality. At the same time, however, a higher number of bank competitors reduces the net benefit from screening prospective borrowers, with the consequence that beyond a certain threshold number of banks, screening yields no net benefit. Re-optimization by banks in light of this fact yields less lending at a higher loan rate. The rise in the loan rate reduces the incentive for borrowers to repay their loans, which implies a reduction in loan quality. The implication is an inverse-U-shaped relationship between the number of banks and loan quality.

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 individual bank loan monitoring has emerged only in a single study, Mester et al. (2007), which is discussed later in this section.

Evidence from Announcement Effects

An examination of whether banks perform 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) and Billett et al. (1995) construct and examine a sample of nearly 1500 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 1000 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. Akhigbe and McNulty (2011) utilize proxy measures of monitoring based on this variable to provide evidence that monitoring improves banks' profit efficiency.

Whereas the above studies study loan announcement effects, Marshall et al. (2014) examine data on forced CEO turnovers, which they argue replaces weaker bank managers with stronger ones that are better monitors of the performances of firm borrowers. They find evidence of significantly positive stock-price effects following forced CEO turnovers at banks than at other firms, which they suggest is consistent with this interpretation.

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 bank-borrowing relationships have yield spreads about 68 basis points lower than bonds issued by firms without such relationships. This finding, they 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. 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 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. 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 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 that borrowers use credit relationships with higher-quality, lower-loss banks to signal their creditworthiness to other parties.

Gorton and Pennacchi (1995) study features of 872 loans sold within a sample of major U.S. banks in 1987 and 1988. 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 characteristics of banks that engaged in loan sales. 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% 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.

Li et al. (2015) seek to assess loan-monitoring gains by relationship lenders by comparing their loan performances during the 2000–2009 interval with those experienced by transactional lenders that specialize in syndicating tradable loans. They find that loans extended by relationship lenders exhibited better performances and less severe quality deterioration, which suggests a positive effect of monitoring on overall loan quality.

Direct Evidence of Bank Monitoring Activities

The studies discussed above provide only indirect evidence regarding the hypothesis that individual banks uniquely provide services as monitors that improve loan quality and, market value. The best sources of data regarding resources that banks apply to the task of monitoring are banks themselves. Such data are typically proprietary, however. Lack of access to such data has limited direct study of the scope of banks' monitoring activities.

The main exception is a study by Mester et al. (2007), which examines data from a single Canadian bank. Mester et al. find that monitoring of transactions deposits held with the lender by borrowers via item-by-item reconciliations of transactions enables verifying reliability of borrowers' statements about flows of accounts receivables and inventories. The analysis of Mester et al. indicates that that provision of transactions deposits 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 recent interesting study by Gustafson et al. (2016) also shines direct light on bank loan monitoring. This interesting study examines evidence from regular monitoring by lead banks in syndicated loans, typically as required by contracts with other loan stakeholders.

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

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

where the parameters α and θ are nonnegative constants governing the magnitude of quadratic resource costs of lending and servicing deposits, and where $\left(\frac{1}{1-i}\right)\frac{c}{2}\left(L^i\right)^2$ is the bank's marginal loan-monitoring cost—*if* the bank chooses to monitor, as discussed below—where c is a positive constant, with $c \leq \alpha$, so that monitoring costs at the median bank ($i = ½$), 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.

The function $\left(\frac{i}{1-i}\right)\frac{c}{2}\left(L^i\right)^2$ implies rising *external* marginal monitoring costs. The derivative of the assumed monitoring cost function with respect to i is $\left(\frac{1}{(1-i)^2}\right)\frac{c}{2}\left(L^i\right)^2 > 0$. 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 $\left(\frac{i}{1-i}\right)\frac{c}{2}\left(L^i\right)^2$ is equal to $\widehat{R}_L^M = \eta r_L$. 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 $\widehat{R}_L^{NM} = (\eta - \delta)r_L$.

Loan Market Equilibrium and Equilibrium Monitoring

The profit-maximizing lending by a monitoring bank is

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

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

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

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 about monitoring when profits are equalized. Substituting optimal quantities of loans as well as securities and deposits 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. 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 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

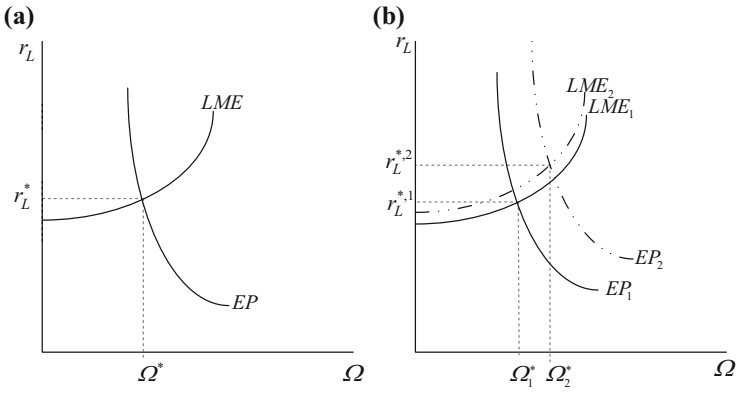

Fig. 6.1 Joint determination of the market loan rate and the equilibrium share of monitoring banks

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 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. It follows that regulations aimed at enhancing the safety and soundness of the banking industry will exert dual effects, as discussed in detail in Chap. 7.

The Relationship Between Banking Competition and Risks

Does bank entry regulation that makes banking markets more concentrated—and potentially less competitive—reduce 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 Chap. 2, 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 and 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 trigger runs is deposit rate ceilings such as those put into place in the United States by the 1933 Glass-Steagall Act. 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.

Arping (2014) argues that whether changes in deposit-market competition induce banks to alter loan pricing in ways that boost exposures to credit risk depends on the degree of loan-rate-setting power they possess. In Arping's model, increased deposit-market competition naturally pushes up the market deposit rate. Less competitive lenders pass along these higher expenses by raising loan rates and consequently reducing exposure to credit risk. More competitive lenders, in contrast, absorb the higher credit risks that result from leaving loan rates at competitive levels when deposit rates rise, which drives down the spread between the two rates. Evidence provided by Delis and Kouretas (2011) from euro-area banks over the 2001–2008 interval is suggestive of the possibility that to the extent that banks in competitive loan markets are unable to boost loan rates across risk classes when open market rates decline, banks may tend to respond to a reduced retail spread by taking on loans with higher risk premiums, thereby expanding the riskiness of their asset portfolios. This line of causation also potentially is consistent with Foos et al.'s (2010) finding of a positive relationship between loan growth and bank riskiness in 16 advanced nations between 1997 and 2007.

Craig and Dinger (2013), focus on interactions between bank deposit markets and wholesale funds markets that can generate higher overall bank funding expenses when competition rises. Craig and Dinger examine data on three risk measures (volatility of return on assets, nonperforming loans, and stock price variability) for 589 banks in 164 metropolitan statistical areas between 1996 and 2007. After taking into account effects that operate jointly across deposit and wholesale-funds markets, they conclude that increased deposit market competition has a net positive effect on bank risk.

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 an 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 counterarguments can be advanced, leading to the opposite conclusion that a freer market is more likely to provide a stable banking environment. Separately, 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 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 Chap. 3, 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 work by Ariss (2010), 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 is more stable in nations with more concentrated banking systems. Mirzei et al. (2013) reach the same essential conclusion in a study of data from more than 1900 banks in 40 nations over the 1999–2008 interval. So do Nguyen et al. (2012), based on

analysis of South Asian banks between 1998 and 2008, except that these researchers find that greater stability for banks with more market power is enhanced when banks diversify across activities that generate both interest and non-interest incomes.

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, Martinez-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 Stebacka (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 large institutions more safe and sound? With respect to size alone, De Haan and Poghosyan (2012) find a negative relationship between bank size and quarterly earnings variability among all U.S. banking institutions between 2004 and 2009. De Haan and Pohosyan find that large banks experienced lower earnings volatility than small banks even after controlling for the degree of market concentration. Nevertheless, they find that greater concentration tends to reduce the effect of bank size on the variability of earnings.

Concerning the market power-risk relationship, Agoraki et al. (2011) report evidence from analysis of data from banks in central and eastern Europe between 1998 and 2005. They find evidence of lower credit risks and reduced likelihood of default for banks possessing greater degrees of market power. In an analysis of 83 nations between 1998 and 2006, Caprio et al. (2014) report results indicating that increased bank concentration was, among other factors, associated with a reduced risk of suffering a crisis in 2008. Fu et al. (2014) utilize the *Z*-score index measure of risk, which encompasses leverage, return on assets, and variability of earnings, to evaluate the relationship between bank pricing power and risk in 14 Asian Pacific nations between 2003 and 2010. They find evidence of an unambiguous negative relationship between pricing power and risk. Leroy and Lucotte (2016) find evidence that increased competition is associated with greater institutional-level risk but nonetheless find that greater competition simultaneously generates less herding behavior that reduces system-wide risks. Berger et al. (2009), who examine data 1999–2005 data from more than 8000 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 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 2500 U.S. banks in 2003 and a panel of about 2600 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 Nicolo (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.

Bhagat et al. (2015) utilize the *Z*-score risk measure to explore the relationship between financial institutions' size and risk between 2002 and 2012. The authors consider data for almost 600 banks, 60 investment banks, and 43 life insurers. They find evidence of a positive relationship between financial-institution size and risk-taking, particularly through increased leverage prior to the 2008 crisis. Delis et al. (2014) likewise find evidence of greater risk among larger banks in an analysis of U.S. data for the 1985–2012 interval. Vallascas and Kaesey (2012) also examine a sample of 600 banks, but their data are from 17 European nations over the 1992–2008 period. They also find a positive relationship between bank size and risk exposures.

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 and 2004 also yield a positive relationship between bank concentration and risk of failure. Agoraki et al. (2011) reach a similar conclusion based on a study of 546 central European banks over the 1994–2005 period, as do Soedarmono et al. (2013) in an analysis of more than 600 banks in 11 emerging Asian nations during the 1994–2009 interval. Furthermore, Carlson and Mitchener (2006) conclude from study of U.S. bank branching data from the 1920s and 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. (2009) 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 (2010) examine 1995–2005 data covering more than 8900 U.S. banks and more than 3600 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. In an examination of European cooperative banks between 1998 and 2009, Fiordelisi and Mare (2014) find that reduced heterogeneity and competition among the banks is positively related to their stability.

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. Beck et al. (2013) further conclude, in the context of a study of international banking data over the 1994–2009 interval, that even though on average greater bank market power is associated with greater bank stability, this relationship is subject to considerable cross-country variation. Beck et al. report evidence indicating that an increased degree of competition is more

likely to be destabilizing if, among other things, nations additionally provide more generous deposit insurance and impose tougher restrictions on allowable activities.

Friexas and Ma (2014) argue that there are strong theoretical grounds for an ambiguous relationship between bank competition and stability because of contrasting effects operating through banks' leverage choices. Consistent with this theoretical ambiguity, in an examination of 1970–2009 data from 160 nations Bretschger et al. (2012) find evidence of separate concentration-stability and concentration-fragility channels, with net effects dependent on relative strengths of those channels. The evidence of the existence of the channels is stronger for developed nations, however, than for less developed countries.

Therefore, currently there is no firm empirical support for the traditional presumption that large banks with greater market power are more stable institutions. Indeed, along at least certain dimensions, considerable evidence points in opposing directions with respect to both size and market power. A meta-analysis of 31 empirical studies of the competition-stability relationship by Zigraiova and Havranek (2016) concludes that the evidence of an overall relationship is, on net, negligible.

Deposit Insurance, "Too Big to Fail" Doctrine, and the Basel Standards

Significantly complicating factors in any assessment of relationships among bank market competition, concentration, and stability are the structure of deposit insurance and 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 $5000) was about 0.25%, 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%. Another 54 years would pass before the FDIC would significantly adjust this premium rate, despite gradual increases in coverage from a $5000 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, the introduction of non-risk-based deposit insurance generally expands risk-shifting incentives. See Demirgüç-Kunt et al. (2008) for a full overview of these and other issues. Work by DeLong and Saunders (2011) has found evidence of adverse risk effects of the introduction of U.S. deposit insurance all but a subset of strongly performing banks among 60 financial institutions in the 1930s. Lambert et al. (2017) find that the 2008 expansion of traditional deposit insurance coverage from $100,000 to $250,000 per deposit induced the institutions most affected to take on riskier loan portfolios. Ioannidou and Penas (2010) likewise find evidence of increased risk-taking by Bolivian banks following introduction of deposit insurance in that nation. Karras et al. (2013) document risk-shifting behavior on the part of Russian depositors as well in response to the advent of deposit insurance.

Duran and Lozano-Vivas (2015) likewise find evidence of risk-shifting effects of deposit insurance for European Union banks during the 2002–2009 interval. Anginer et al. (2014) provide evidence from analysis of more than 4000 banks in 96 nations between 2004 and 2009 that a predominant stabilization effect of deposit insurance exists during period of banking turbulence but that moral-hazard risk-shifting incentives predominate during good times. Angkinand and Wihlborg (2010) show that the nature of relationship between deposit insurance coverage and banks' risk-shifting incentives depends on several factors, including whether government or foreign ownership may be involved and on the nature of shareholder rights and characteristics of governance. Based on a study of data from emerging-market nations, Angkinand and Wihlborg find evidence of a U-shaped relationship between deposit insurance coverage and risk-taking incentives.

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%, 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% 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% of insured deposits, and since the late 2000s most banks have paid deposit insurance premium rates ranging from 0.05 to 0.12%. 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—the scope of which Demirgüc-Kunt et al. (2015) show has generally expanded since the late-2000s crisis—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 2-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—see Kaufman (2014), Deng et al. (2007), and Ennis and Malek (2005)—continues to offer competitive advantages over smaller institutions. Based on data from the 1991–2004 period, Brewer and Jagtiani (2013) estimate that acquirers in eight bank mergers paid at least $15 billion in premiums to boost the combined entities into the too-big-too-fail range established by U.S. regulators. Gandhi and Lutig (2015) estimate that consequent past average annual per-bank too-big-to-fail savings in equity expenses for the largest U.S. banks have amounted to about $2.76 billion.

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, as documented by Barth et al. (2012) and Gandhi and Lustig (2015). Echoing results obtained from earlier banking data by Hughes and Mester (1993), Jacewitz and Pogach (2014) utilize panel data on deposit rate observations for money market deposit accounts at more than 35,000 U.S. bank branches to demonstrate that the largest banks derive a competitive advantage from their too-big-to-fail status. Separately, Molyneux et al. (2014) find in an analysis of bank assistance programs conducted by twelve nations in the wake of the global financial meltdown that banks' efforts to boost scale sufficiently to become deemed too big to fail does indeed increase their likelihood of becoming rescue beneficiaries.

Gorton and Tallman (2016) draw on historical analysis of policies of pre-central-bank private clearinghouses to rationalize a too-big-to-fail policy as a" reasonable response to the threat posed to large banks by the vulnerability of short-term debt to

runs." An important distinction today, of course, is the public rather than private guarantees that now are provided to too-big-to-fail banks. The solvency guarantees provided to large banks by governments substantially dwarf the meager funding pools established by private clearinghouses in years past. Marshall et al. (2016) estimate that when combined with governmental deposit-insurance promises, the total amount of guarantees offered by the U.S. government amount to more than 60% of the assets of the entire financial sector.

Demirgüc-Kunt and Huizinga (2013) suggest that banks derive unambiguous benefits from a too-big-to-fail policy only when the fiscal health of national governments is strong. Shocks from which a too-big-to-fail policy might help shield larger banks can work against this outcome if such disturbances weaken the capability of governments to bail out their banking systems. Consequently, for banks in nations with weaker fiscal positions, downsizing can be the appropriate response to such shocks.

As pointed out by Jones and Oshinsky (2009), pursuit of a too-big-to-fail policy impinges on the strength of governmental deposit insurance guarantees. Jones and Oshinsky document that increased concentration of deposits within larger banks substantially increased insolvency risk to the FDIC's insurance fund during the years leading up the global financial crisis, during which the FDIC did indeed exhaust its resources and find itself dependent upon loans from the U.S. Treasury.

Furthermore, as noted by Strahan (2008), small U.S. banks benefit from various explicit and implicit government and Federal Reserve subsidies. Nevertheless, not being too big to 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.

In contrast, 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.

DeYoung et al. (2013) reframe the problem in terms of a "too-complex-to-fail" issue. Within a model of a repeated game between a government resolution authority and the banking system, DeYoung et al. conclude that limits on technologies for resolutions of failures and political-economy incentives create a bias in favor of bailing out more complex institutions—which in turn gives institutions incentives to become more complex.

Capital Regulation and the Three Pillars of the Basel Framework

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% 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% 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 various versions of so-called Basel system around three so-called 'pillars': Risk-based capital requirements, supervisory discipline, and increased market discipline [see, for instance, Tarullo (2008)]. Issues related to the various Basel requirements are considered in remaining chapters.

References

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.

Acharya, Viral, and Tanju Yorulmazer. 2008. Information contagion and bank herding. *Journal of Money, Credit, and Banking* 50: 215–231.

Agoraki, Mari-Eleni, Manthos Delis, and Fotios Pasiouras. 2011. Regulations, competition, and bank risk-taking in transition countries. *Journal of Financial Stability* 7: 38–48.

Akhigbe, Aigbe, and James McNulty. 2011. Bank monitoring, profit efficiency, and the commercial lending business model. *Journal of Economics and Business* 63: 531–551.

Allen, Franklin, and Douglas Gale. 2004. Competition and financial stability. *Journal of Money, Credit, and Banking* 36: 453–480.

Anderlini, Luca. 1986. Central banks and moral hazard. Economic Theory Discussion Paper No. 103, Department of Applied Economics, University of Cambridge.

Anginer, Deniz, Asli Demirgüç-Kunt, and Min Zhu. 2014. How does deposit insurance affect bank risk? Evidence from the recent crisis. *Journal of Banking and Finance* 48: 312–321.

Angkinand, Apanard. 2009. Banking regulation and the output cost of banking crises. *Journal of International Financial Markets, Institutions, and Money* 19: 240–257.

Angkinand, Apanard, and Clas Wihlborg. 2010. Deposit insurance coverage, ownership, and banks' risk-taking in emerging markets. *Journal of International Money and Finance* 29: 252–274.

Ariss, Rima Turk. 2010. On the implications of market power in banking: Evidence from developing countries. *Journal of Banking and Finance* 34: 765–775.

Arping, Stefan. 2014. Bank competition, welfare, and stability: The interplay between loan and deposit markets. Unpublished Manuscript, University of Amsterdam, December.

Barth, James, Apanard Prbha, and Philip Swagel. 2012. Just how big is the too big to fail problem? Milken Institute Current Views, March.

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.

Beck, Thorsten, Olivier De Jonghe, and Glenn Schepens. 2013. Bank competition and stability: Cross-country heterogeneity. *Journal of Financial Intermediation* 27: 218–244.

Berger, Allen, Leora Klapper, and Rima Turk-Ariss. 2009. Bank competition and financial stability. *Journal of Fianncial Services Research* 2009: 99–118.

Bhagat, Sanjai, Brian Bolton, and Jun Lu. 2015. Size, leverage, and risk-taking of financial institutions. *Journal of Banking and Finance* 59: 520–537.

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.

Boone, Jan. 2001. Intensity of competition and the incentive to innovate. *International Journal of Industrial Organization* 19: 705–726.

Boyd, John, and Gianni De Nicolo. 2005. The theory of bank risk taking and competition revisited. *Journal of Finance* 60: 1329–1343.

Boyd, John, and Mark Gertler. 1993. U.S. commercial banks: Trends, cycles, and policy. In *NBER Macroeconomics Annual*, ed. Olivier Blanchard and Stanley Fischer, 319–368. Cambridge, MA: MIT Press.

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.

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.

Bretschger, Lucas, Vivien Kappel, and Therese Werner. 2012. Market concentration and the likelihoods of financial crises. *Journal of Banking and Finance* 36: 3336–3345.

Brewer, Elijah III, and Julapa Jabtiani. 2013. How much did banks pay to become too big to fail and to become systemically important? *Journal of Financial Services Research* 43: 1–35.

Calomiris, Charles, and Charles Kahn. 1991. The role of demandable debt in structuring optimal banking arrangements. *American Economic Review* 81: 497–513.

Caminal, Ramon, and Carmen Matutes. 2002. Market power and banking failures. *International Journal of Industrial Organization* 20: 1341–1361.

Caprio, Gerard Jr., Vencenzo D'Apice, Giovanni Ferri, and Giovanni Walter Pupolo. 2014. Macro-financial determinants of the great financial crisis: Implications for financial regulation. *Journal of Banking and Finance* 44: 114–129.

Carletti, Elena. 2008. Competition and regulation in banking. In *Handbook of Financial Intermediation and Banking*, ed. Anjan Thakor and Arnoud Boot, 449–482. Amsterdam: Elsevier.

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, 2002.

Carletti, Elena, Vittoria Cerasi, and Sonja Daltung. 2007. Multiple-bank lending: Diversification and free-riding in monitoring. *Journal of Financial Intermediation* 16: 425–451.

Carlson, Mark, and Kris James Mitchener. 2006. Branch banking, bank competition, and financial stability. *Journal of Money, Credit, and Banking* 38: 1293–1328.

Cole, Rebel, and Lawrence White. 2012. Déjà vu all over again: The causes of U.S. commercial bank failures *this* time around. *Journal of Financial Services Research* 42: 5–29.

Coleman, Anthony, Neil Esho, and Ian Sharpe. 2006. Does bank monitoring influence loan contract terms. *Journal of Financial Services Research* 30: 177–198.

Craig, Ben, and Valeriya Dinger. 2013. Deposit market competition, wholesale funding, and bank risk. *Journal of Banking and Finance* 37: 3605–3622.

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.

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.

Datta, Sudip, Mai Iskandar-Datta, and Ajay Patel. 1999. Bank monitoring and the pricing of corporate debt. *Journal of Financial Economics* 51: 435–449.

De Haan, Jakob, and Tigran Poghosyan. 2012. Bank size, market concentration, and bank earnings volatility in the United States. *Journal of International Financial Markets, Institutions, and Money* 22: 35–54.

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 Vries, C.G. 2005. The simple economics of bank fragility. *Journal of Banking and Finance* 29: 803–825.

Delis, Manthos, and Georgios Kouretas. 2011. Interest rates and bank risk-taking. *Journal of Banking and Finance* 35: 840–855.

Delis, Manthos, Iftekhar Hasan, and Efthymios Tsionas. 2014. The risk of financial intermediaries. *Journal of Banking and Finance* 44: 1–12.

DeLong, Gayle, and Anthony Saunders. 2011. Did the introduction of fixed-rate federal deposit insurance increase long-term bank risk-taking? *Journal of Financial Stability* 7: 19–25.

Demirgüç-Kunt, Asli, and Harry Huizinga. 2013. Are banks too big to fail or too big to save? International evidence from equity prices and CDS spreads. *Journal of Banking and Finance* 37: 875–894.

Demirgüç-Kunt, Asli, Edward Kane, and Luc Laeven, eds. 2008. *Deposit Insurance Around the World: Issues of Design and Implementation*. Cambridge, MA: MIT Press.

Demirgüc-Kunt, Asli, Edward Kane, and Luc Laeven. 2015. Deposit insurance around the world: A comprehensive analysis and database. *Journal of Financial Stability* 20: 155–183.

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.

DeYoung, Robert, Michal Kowalik, and Jack Reidhill. 2013. A theory of failed bank resolution: Technological change and political economics. *Journal of Financial Stability* 9: 612–627.

Diamond, Douglas. 1984. Financial intermediation and delegated monitoring. *Review of Economic Studies* 51: 393–414.

Diamond, Douglas. 1991. Monitoring and regulation: The choice between bank loans and directly placed debt. *Journal of Political Economy* 99: 689–721.

Diamond, Douglas. 1996. Financial intermediation as delegated monitoring: A simple example. *Federal Reserve Bank of Richmond Economic Quarterly* 82 (3): 51–66.

Diamond, Douglas, and Phillip Dybvig. 1983. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy* 91: 401–419.

Diamond, Douglas, and Raghuram Rajan. 2000. A theory of bank capital. *Journal of Finance* 55: 2431–2465.

Diamond, Douglas, and Raghuram 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. 2006. Money in a theory of banking. *American Economic Review* 96: 30–53.

Dowd, Kevin. 1992. Models of banking instability: A partial review of the literature. *Journal of Economic Surveys* 6: 107–132.

Dowd, Kevin. 1994. Competitive banking, bankers' clubs, and bank regulation. *Journal of Money, Credit, and Banking* 26: 289–308.

Dowd, Kevin. 2000. Bank capital adequacy versus deposit insurance. *Journal of Financial Services Research* 17: 7–15.

Duran, Miguel, and Ana Lozano-Vivas. 2015. Moral hazard and the financial structure of banks. *Journal of International Financial Markets, Institutions, and Money* 34: 28–40.

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.

Ennis, Huberto, and H.S. Malek. 2005. Bank risk of failure and the too-big-to-fail policy. *Federal Reserve Bank of Richmond Economic Quarterly* 91 (Spring): 21–44.

Fiordelisi, Franco, and Davide Salvatore Mare. 2014. Competition and financial stability in European cooperative banks. *Journal of International Money and Finance* 45: 1–16.

Foos, Daniel, Lars Norden, and Martin Weber. 2010. Loan growth and riskiness of banks. *Journal of Banking and Finance* 34: 2929–2940.

Freeman, Scott. 1988. Banking as the provision of liquidity. *Journal of Business* 61: 45–64.

Freixas, Xavier, and Kebin Ma. 2014. Banking competition and stability: The role of leverage. Tilburg University Discussion Paper No. 2014-048, August 29.

Fu, Xiaoqing (Maggie), Yongjia (Rebecca) Lin, and Philip Molyneux. 2014. Bank competition and financial stability in Asia Pacific. *Journal of Banking and Finance* 38: 64–77.

Gande, Amar, Manu Puri, Anthony Saunders, and Ingo Walter. 1997. Bank underwriting of debt securities: Modern evidence. *Review of Financial Studies* 10: 1175–1202.

Gandhi, Priyank, and Hanno Lustig. 2015. Size anomalies in U.S. bank stock returns. *Journal of Finance* 70: 733–768.

Goedde-Menke, Michael, Thomas Langer, and Andreas Pfingsten. 2014. Impact of the financial crisis on bank run risk—Danger of the days after. *Journal of Banking and Finance* 40: 522–533.

Goldstein, L., and A. Pauzner. 2005. Demand deposit contracts and the probability of bank runs. *Journal of Finance* 60: 1293–1328.

Gomez, Fabianna, and Jorge Ponce. 2014. Bank competition and loan quality. *Journal of Financial Services Research* 46: 215–233.

Goodhart, Charles. 1988. *The Evolution of Central Banks*. Cambridge, MA: MIT Press.

Gorton, Gary, and George Pennacchi. 1995. Banks and loan sales: Marketing nonmaketable assets. *Journal of Monetary Economics* 35: 389–411.

Gorton, Gary, and Ellis Tallman. 2016. Too-bit-to-fail before the Fed. Federal Reserve Bank of Cleveland Working Paper 16-12, May.

Gustafson, Matthew, Ivan Ivanov, and Ralf Meisenzahl. 2016. Bank monitoring: Evidence from syndicated loans. Unpublished Working Paper, Pennsylvania State University and Board of Governors of the Federal Reserve System, August.

Hainz, Christa, Laurent Weill, and Christophe Godlewski. 2013. Bank competition and collateral: Theory and evidence. *Journal of Financial Services Research* 44: 131–148.

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.

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.

Hughes, Joseph, and Loretta Mester. 1993. A quality and risk-adjusted cost function for bank: Evidence on the 'too-big-to-fail' doctrine. *Journal of Productivity Analysis* 4: 293–315.

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.

Ioannidou, Vasso, and Maria Fabiana Penas. 2010. Deposit insurance and bank risk-taking: Evidence from internal loan ratings. *Journal of Financial Intermediation* 19: 95–115.

Iyer, Rajkamal, and Manju Puri. 2008. Understanding bank runs: The importance of depositor-bank relationships and networks. NBER Working Paper 14280, August.

Jacewitz, Stefan, and Jonathan Pogach. 2014. Deposit rate advantages at the largest banks. FDIC Center for Financial Research Working Paper 2014-02, February.

Jacklin, C. 1987. Demand deposits, trading restrictions, and risk sharing. In *Contractual Arrangements for Intertemporal Trade*, ed. Edward Prescott and Neil Wallace. Minneapolis, MN: University of Minnesota Press.

James, Christopher. 1987. Some evidence on the uniqueness of bank loans. *Journal of Financial Economics* 19: 217–235.

Jones, Kenneth, and Robert Oshinsky. 2009. The effect of industry consolidation and deposit insurance reform on the resiliency of the U.S. bank insurance fund. *Journal of Financial Stability* 9: 57–88.

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.

Karas, Alexei, William Pyle, and Koen Schoors. 2013. Deposit insurance, banking crises, and market discipline: Evidence from a natural experiment on deposit flows and rates. *Journal of Money, Credit, and Banking* 45: 179–200.

Kaufman, George. 2014. Too big to fail in banking: What does it mean? *Journal of Financial Stability* 13: 214–223.

Keeley, Michael. 1990. Deposit insurance, risk, and market power in banking. *American Economic Review* 80: 1183–1200.

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.

Kiss, Hubert, Ismael Rodriguez-Lara, and Alfonso Rosa-Garcia. 2012. On the effects of deposit insurance and observability on bank runs: An experimental study. *Journal of Money, Credit, and Banking* 44: 1651–1665.

Kopecky, Kenneth, and David VanHoose. 2006. Capital regulation, heterogeneous monitoring costs, and aggregate loan quality. *Journal of Banking and Finance* 30: 2235–2255.

Koskela, Erkki, and Rune Stenbacka. 2000. Is there a tradeoff between bank competition and financial fragility? *Journal of Banking and Finance* 24: 1853–1873.

Kwan, Simon. 2001. Financial modernization and banking theories. Federal Reserve Bank of San Francisco Economic Letter, No. 2001-37, December 21.

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.

Lambert, Claudia, Felix Noth, and Ulrich Schüwer. 2017. How do insured deposits affect bank risk? Evidence from the Emergency Economic Stabilization Act. Journal of Financial Intermediation.

Lee, Kwang-Won, and Ian Sharpe. 2009. Does a bank's loan screening and monitoring matter? *Journal of Financial Services Research* 35: 33–52.

Leroy, Aurelien, and Yannick Lucotte. 2016. Is there a competition-stability trade-off in European banking? *Journal of Financial Markets, Institutions, and Money*.

Li, Yutao, Anthony Saunders, and Pei Shao. 2015. The monitoring incentive of transactional and relationship lenders: Evidence from the syndicated loan market. *Journal of Money, Credit, and Banking* 47: 702–735.

Marinč, Matej. 2008. Essays on Bank Monitoring, Regulation, and Competition, PhD. Dissertation, University of Amsterdam.

Marshall, Andrew, Laura McCann, and Patrick McColgan. 2014. Do banks really monitor? Evidence from CEO succession decisions. *Journal of Banking and Finance* 46: 118–131.

Marshall, Liz, Sabrina Pellerin, and John Walter. 2016. Bailout barometer: How large is the financial safety net? *Federal Reserve Bank of Richmond*, February 3. https://www.richmondfed.org/safetynet/

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.

Matutes, Carmen, and Xavier Vives. 1996. Competition for deposits, fragility, and insurance. *Journal of Financial Intermediation* 5: 184–216.

Matutes, Carmen, and Xavier Vives. 2000. Imperfect competition, risk taking, and regulation in banking. *European Economic Review* 44: 1–34.

Mester, Loretta, Leonard Nakamura, and Micheline Renault. 2007. Transactions accounts and loan monitoring. *Review of Financial Studies* 20: 529–556.

Mirzei, Ali, Tomoe Moore, and Guy Liu. 2013. Does market structure matter on banks' profitability and stability? Emerging vs. advanced economies. *Journal of Banking and Finance* 37: 2920–2937.

Molyneux, Philip, Klaus Schaeck, and Tim Mi Zhou. 2014. 'Too systemically important to fail' in banking—Evidence from bank mergers and acquisitions. *Journal of International Money and Finance* 49: 258–282.

Nguyen, My, Michael Skully, and Shrimal Perera. 2012. Market power, revenue diversification, and bank stability: Evidence from selected South Asian countries. *Journal of International Money and Finance* 22: 897–912.

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.

Ogura, Yoshiaki. 2006. Learning from a rival bank and lending boom. *Journal of Financial Intermediation* 15: 535–555.

Panzar, J.C., and J.N. Rosse. 1987. Testing for 'monopoly' equilibrium. *Journal of Industrial Economics* 35: 443–456.

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.

Schaeck, Klaus, and Martin Čihák. 2010. Competition, efficiency, and soundness in banking: An industrial organization perspective. European Banking Center Bank Discussion Paper No. 201068S, July.

Schaeck, Klaus, Martin Čihák, and Simon Wolfe. 2009. Are competitive banking systems more stable? *Journal of Money, Credit, and Banking* 41: 711–734.

Schotter, Andrew, and Tanju Yorulmazer. 2009. On the dynamics and severity of bank runs: An experimental study. *Journal of Financial Intermediation* 18: 217–241.

Smith, Bruce. 1984. Private information, deposit interest rates, and the 'stability' of the banking system. *Journal of Monetary Economics* 14: 293–317.

Soedarmono, Wahyoe, Fouad Machrouh, and Amine Tarazi. 2013. Bank competition, crisis, and risk taking: Evidence from emerging markets in Asia. *Journal of International Financial Markets, Institutions, and Money* 23: 196–221.

Stern, Gary, and Ron Feldman. 2004. *Too Big to Fail: The Hazards of Bank Bailouts*. Washington, DC: Brookings Institution.

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 *Handbook of Financial Intermediation and Banking*, ed. Anjan Thakor and Arnoud Boot, 97–132. Amsterdam: Elsevier.

Tarullo, Daniel. 2008. *Banking on Basel: The Future of International Financial Regulation*. Washington, DC: Peterson Institute for International Economics.

Vallascas, Francesco, and Kevin Keasey. 2012. Bank resilience to systemic shocks and the stability of banking systems: Small is beautiful. *Journal of International Money and Finance* 31: 1745–1776.

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.

Wicker, Elmus. 1996. *The Banking Panics of the Great Depression*. Cambridge, UK: Cambridge University Press.

Zigraiova, Diana, and Tomas Havranek. 2016. Bank competition and financial stability: Much ado about nothing? *Journal of Economic Surveys* 30: 944–981.

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 (2007) have provided 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 implementation of Basel capital standards.

© Springer-Verlag GmbH Germany 2017

D. VanHoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-662-54326-9_7

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 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 Chap. 2. 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 consider 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 EF_1 is the efficient frontier associated with alternative 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

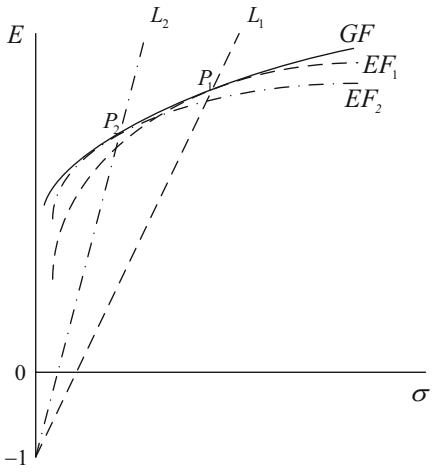

Fig. 7.1 Efficient portfolios and the probabilities of bank insolvency at different capital-asset ratios

return. Blair and Heggestad (1978) show 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 $\bar{k}$ corresponding to the probability of failure given by the square root of the reciprocal of the slope of the ray $\bar{L}$ 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 $\bar{P}$ 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

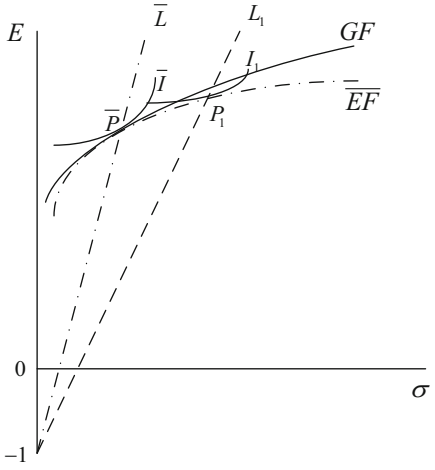

Fig. 7.2 A bank capital requirement and portfolio choice

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, 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.

Indeed, Akhigbe et al. (2012) offer evidence of "capital signaling" that pushed down stock price performances of higher-capitalized banks during the recent global financial crisis. They suggest that investors viewed banks with larger levels of capital as likely as well to hold higher-risk assets. Perotti et al. (2011) argue that access to "tail-risk assets"—assets offering the possibility for realizations of negative return outcomes in excess of any feasible level of capital—could account for such an observed relationship.

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 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 original Basel I system implemented in the early 1990s. 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 and King (1995), and Gordy (2003) for more detailed discussions of linking capital regulation to risk-weighted assets]. Ultimately, of course, the Basel I risk weights were not closely aligned with actual asset risks—see, for instance, Barakova and Palvia (2014).

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 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. Fegatelli (2010) argues, however, that what these authors classify as an option value of deposit insurance actually is a limited liability option already incorporated into the value of bank equity. Gale (2010) additionally concludes that even if the option-value approach is correct, it does not preclude the shifting or expansion of bank risks in response to capital regulation.

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 on portfolio-managing banks? Rochet (1992) suggests that part of the answer depends on whether one assumes complete or incomplete markets. He considers a setting with exogenously fixed capital. 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: (1) no increase in capital but investment in a combination of one risky asset and the riskless asset, (2) complete specialization in a single risky asset and just meeting the requirement, or (3) 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.

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 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 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 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." Numerous researchers have considered the manner in which a bank balances costs and benefits across its balance sheet when subjected to capital regulation.

Admati et al. (2013) and Admati and Hellwig (2014) argue that the private costs that banks must incur to obtain equity and other funds performing capital-like functions are very so low—so low that they are dwarfed by the social benefits of higher capital cushions. As Calomiris (2013) points out, however, the costs that regulatory capital requirements impose on banking firms are not negligible, particularly in the short run, as discussed by Dagher et al. (2016). Also not negligible, as Aiyar et al. (2015) note, are the costs borne by prospective borrowers who do not receive loans as a consequence of capital limitations, actual borrowers who incur higher interest expenses on loans that they do receive, and bank deposits that earn lower returns as a result of capital regulations.

Calomiris (2013) also argues that the relevant issue is not the absolute magnitudes of capital cushions but the "capital structure choices of banks *relative to the risk of their assets*" (Calomiris 2013, p. 19). That is, socially optimal capital standards and associated regulations must account *both* for the social benefits of expanded cushions against ex post losses *and* the decisions made ex ante by banking firms in reaction to the private costs generated by these standards and regulations. Banks' ex ante choices govern the magnitudes of their asset risks and, consequently the scale of the ex post social losses that higher capital requirements are aimed to cushion.

Incentives and Capital Requirements

Modern models of constrained bank balance-sheet adjustments to capital requirements focus on the effects of risk-based capital regulation alongside pure leverage-based standards. Several approaches also seek take into account dynamic elements of the problem that banks confront. Banks' choices that influence their asset risks

depend considerably on the competitive structures of the markets that they confront, however.

Perfect Competition Models of Bank Capital Regulation

Much work evaluating incentive effects of bank capital regulation assume perfectly elastic supplies of assets and demands for liabilities on the part of the public. Estrella (2004a), for instance, develops a framework that, in contrast to portfolio-management models, allows a bank to adjust its liability structure. In Estrella's model, a bank facing elements of all three pillars of the Basel standards 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: one that cannot raise sufficient funds to invest in an asset and closes, one that meets the capital constraint and must issue debt, and one that has sufficient capital to invest without issuing debt and 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 (2004b) 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 Chap. 2. 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. 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 consider a short-run horizon in which equity is fixed at a level $\bar{E}^i$ 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, $\pi^i = r_L L^i + r_S S^i - r_D D^i - r_N N^i - C^i(L^i, S^i, D^i, N^i)$. 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. 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

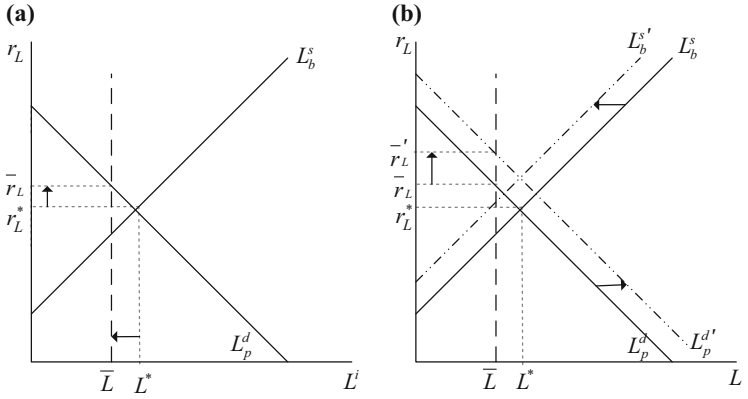

Fig. 7.3 Short-run implications of a binding risk-based capital requirement with representative perfectly competitive banks

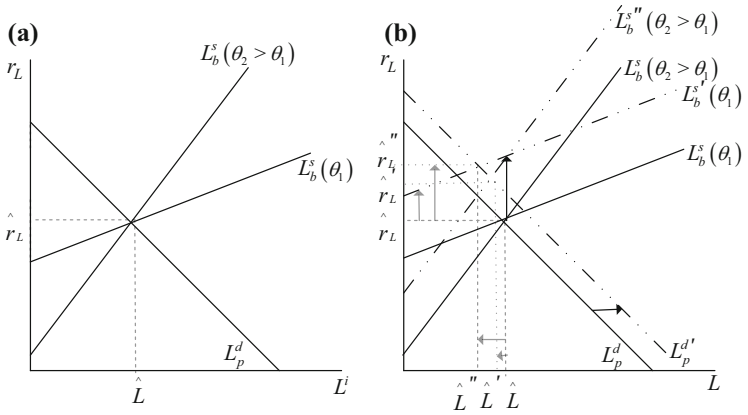

Fig. 7.4 Long-run effects of a binding risk-based capital ratio with representative perfectly competitive banks

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 $\bar{L}$, and the market loan rate rises above the level in an unconstrained equilibrium, at $\bar{r}_L$. Panel (b) exhibits the effects of an increase in the level of the risk-free rate that translates into higher interest rates 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. 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 $\bar{r}_L$ to $\bar{r}_L{}'$.

Panel (a) of Fig. 7.4 shows that in the long run, which Barnea and Kim (2014) and De Jonghe and Öztekin (2015) have estimated to be a period of 4–6 years for most banks but which Berger et al. (2008) and Memmel and Raupach (2010) have found to be less than 3 years for larger institutions, banks constrained by capital regulation must raise additional higher-marginal-cost equity to support given volumes of lending. Their total marginal costs of lending consequently are steeper with higher required capital ratios. An increase in the required capital ratio thereby 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, $\theta_2 > \theta_1$.

As Aiyar et al. (2014a, b) point out, an important limitation of the above analysis—indeed, much of the research surveyed in this chapter—is its closed-economy structure. As discussed in Chap. 4, a large number of banks have multinational operations. In principle, these banks can respond to binding domestic capital constraints by reallocating funds across subsidiaries located in jurisdictions with differing capital standards or degrees of enforcement. These authors find evidence in UK data of the occurrence of such cross-border reallocational responses to capital regulation among multinational banks. We return to the international complications raised by capital standards and other forms of banking regulation in greater detail in Chap. 10.

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 model, banks choose to hold either riskless or risky asset portfolios. Higher expected profits in the second period induce an unregulated bank to cut risk in the first period to lower the probability of failure and 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 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 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) concludes that inclusion of a leverage-ratio requirement—a feature adopted in the current rendition of the Basel standards—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. More recently, Dermine (2015) argues within

the context of a theoretical model of portfolio credit risk that a non-risk-based leverage-ratio standard effectively places a floor on the risk-based capital ratio. His conclusion is that such a policy thereby reduces both credit risk and liquidity risk and thereby reduces the likelihood of a bank run.

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 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 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 et al. (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. When asset and capital returns are endogenous, however, banks' incentives to boost asset risk increase with tougher capital requirements. Hellmann et al. conclude 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 induces 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 Chap. 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 Chap. 2, 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, 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

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) discuss 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, (1) higher bank equity signals to capital markets that bank insiders consider their assets to be riskier; (2) 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 moral hazard problems on the liability side of banks' balance sheets. Diamond and Rajan (2000, 2001) are important contributions in this vein. Building on 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. Their emphasis on financial fragility on the liability side of the bank's balance sheet implies that the demandable nature of deposits is crucial for 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 and 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

Work by Diamond (1984, 1991) and others has emphasized banks' key role in monitoring loan assets for moral hazard risks. 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 choices in the first period from an endowment of risky loan opportunities, and choose disutility-generating effort that boosts 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 et al. (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 2007–2009 bank solvency meltdown, 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, 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 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 effort to allow for diversity in responses to capital regulation is provided by Almazan (2002), who considers an setting in which banks balance capital versus monitoring expertise, 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 Chap. 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

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

where $\left(\frac{i}{1-i}\right)\frac{c}{2}\left(L^i\right)^2$ is bank i's monitoring cost and where $\widehat{R}_L^M = \eta r_L$ for a bank that monitors and $\widehat{R}_L^{NM} = (\eta - \delta) r_L$ 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 $L^i = \hat{L}\, i = \frac{\bar{E}}{\mu}$, where μ again denotes the capital requirement ratio.

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 $\hat{r}_L = \frac{l_0 - \frac{\bar{E}}{\mu}}{l_1}$, 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 $\hat{r}_L = \frac{1}{2}\left(\frac{\Omega}{1-\Omega}\right)\left(\frac{c}{\delta}\right)\left(\frac{\bar{E}}{\mu}\right)$, 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 Chap. 6, under capital 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 use 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 fraction that otherwise would have monitored their loans in the absence of capital regulation [the fraction Ω^* in Fig. 6.1]. 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 further reduction in constrained bank lending reduces marginal monitoring costs, thereby giving some banks a greater incentive to monitor. 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

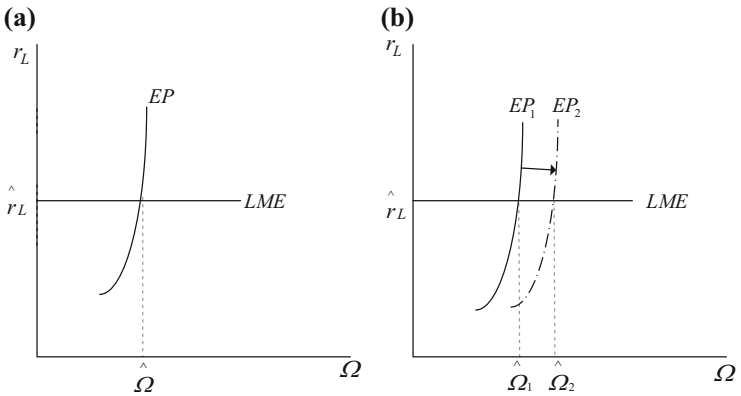

Fig. 7.5 Effects of a binding risk-based capital ratio with heterogeneous competitive banks

monitor their loans for moral hazard risks. As a consequence, aggregate loan quality may improve or worsen. Further analysis by Kopecky and VanHoose (2012) indicates that while in principle a capital requirement might be established to move monitoring incentives into line to achieve a socially optimal degree of loan quality, the informational requirements faced by a regulatory in implementing such a policy would be substantial.

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. 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.

Hakanes and Schnabel (2011) contemplate a heterogenous setting populated with large and small banks. They apply a Hotelling (1929)-style location model to examine the competitive effects of allowing competing large and small banks to choose between fixed-weight versus internal-risk-weight-based capital regulation. Their analysis indicates that under conditions in which small banks maximize profits by choosing a fixed-weight capital requirement while large banks opt for an internal-risk-weight-based approach, larger banks gain a cost advantage. To compete effectively, small banks must respond by selecting riskier portfolios.

There has been relatively little empirical work evaluating potential impacts of capital regulation on the competitive structure of the banking industry. 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 could be placed at a competitive disadvantage, and Lang et al. (2008), who reach the opposite conclusion with respect to large credit card lenders. In addition, Schliephake and Kirstein (2013) argue that by constraining banks' lending activities binding capital requirements constitute exogenous capacity constraints that generate Cournot behavior, which in turn rationalizes applying Cournot-based oligopoly analysis to the behavior of capital-constrained banks.

In recent work, Dia and VanHoose (2017) develop a Cournot-oligopoly model in which banks face fixed costs and can enter or exit the industry to the point at which zero economic profits are attained. The model is applied to examine, both analytically and via calibrated simulations, the effects of the recent Basel III-style tightening of capital standards combined with higher fixed costs to banks of complying with regulations. One key implication of the analysis is that both toughened capital regulation and higher regulatory compliance costs that boost banks' fixed operating costs generate a reduction in the equilibrium quantity of loans, an increase in the market loan rate, and a reduction in the number banks. Hence, bank loan markets become more concentrated, an effect that increases in magnitude with a larger value of the market loan demand elasticity. Another key implication is that these effects also alter the incentives for banks to monitor their loans and, as in the Kopecky and VanHoose (2006) model, eliminate sources of loan losses. The analysis and simulations that Dia and VanHoose conduct indicate that a tighter capital constraint and higher regulatory compliance costs reduce the incentives for banks to monitor their loans. Thus, the composition of the reduced market volume of loans shifts to unmonitored loans subject to losses, which reduces aggregate loan quality.

Capital, Shocks and Procyclicality, and Bank Performance

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 original 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 during that nation's 1989–1996 period of financial distress. They find evidence that Japan's banks were capital constrained and 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 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. In a study of quarterly dat for U.S. bank holding companies during the 2006–2011 interval, however, Leung et al. (2015) find evidence of the opposite relationship.

Beatty and Gron (2001) examine data for 438 publicly traded U.S. bank holding companies between 1986 and 1995. 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. 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.

Brewer et al. (2008) examine 1992–2005 data from the world's largest 78 banks in an effort to evaluate determinants of their capital positions. They find that that observed capital ratios are indeed higher in nations with tougher capital regulations. Other important factors include the relative size of a nation's banking system, to which bank capitalization is inversely related, and the degree to which nations' regulators engage in prompt corrective action, to which bank capitalization is positively related.

In an analysis of 1999–2005 data from more than 2600 financial institutions in 10 European countries, Schaeck and Cihák (2012) add to this list of capital-ratio-boosting factors a greater reliance on relationship-based lending, stronger shareholder rights, and use of deposit-insurance systems that exclude non-deposit creditors. Schaeck and Cihák also find that higher capital ratios also are associated with more competitive banking markets. They argue that this result supports Allen et al.'s (2011) hypothesis that market discipline made possible by more competitive markets benefits borrowers by inducing banks to commit costly capital to loan monitoring than reducing credit supply and bringing about higher market loan rates.

In a creative study, Furfine (2001) develops an intertemporal model that yields optimizing conditions that guide bank decision-making with and without capital regulation. He utilizes data from FDIC call reports for 362 banks with assets exceeding $1 billion to estimate these optimizing conditions directly. Furfine then uses 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 on 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 U.S. "credit crunch" of 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 a panel of 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.

In contrast, a matched-bank study by Carlson et al. (2013) over the 2001–2011 interval finds negligible effects of capital ratios on bank lending over most of the period and particularly during years of economic expansion. Only over the buildup to and during the global financial crisis do the authors find a significant response of bank loans to capital ratios.

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 capital necessarily reduces 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, and some evidence—see, for instance Guidara et al. (2013)—indicates that capital buffers vary procyclically. Hence, banking inherently tends to be a procylical industry, as verified by Jokivuolle et al. (2015) in a study of 1982–2012 data from nine European nations and by Shim (2013) in analysis of 1992–2011 U.S. data. As noted by Goodhart et al. (2004), 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.

Goodhart et al., note that regulation adds to banking's procylical tendencies. The bank supervisory process tends to press banks to constrict lending during contractions 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.

As examined by Drummond (2009), Repullo (2013), and Athanasoglou et al. (2014), a number of observers have suggested that the Basel systems of risk-based capital requirements enhance the procyclical effects of bank regulation. Indeed, under the Basel standards, minimum capital changes 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 the Basel standards—which he suggests could help mitigate the extent to which capital actually responds to downturn-induced boosts in risks. Nevertheless, in the absence of countercyclical adjustments such as recommended by Jokivuolle et al. (2014), 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 Basel regulatory 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? 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 through 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 Basel 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. They conclude that enforcement of the Basel regulatory system has the potential to considerably add to banking procyclicality. Simulations conducted by Anderson (2011) yield similar conclusions. Agénor and da Silva (2012) reach this same conclusion from a theoretical analysis of the effects of alternative macroeconomic shocks under different forms of capital regulation. Notably, they find that regulatory capital buffers—see, for instance, Drehmann et al. (2010)—intended to induce countercyclical movements actually have the potential to bring about exactly the opposite effects.

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 et al. (2004) use 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 2600 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 requirements based on ex ante *anticipations* of asset risk changes instead of ex post *observations* of altered risks. Although 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 would considerably smooth business-cycle peaks and troughs.

Brei and Gambacorta (2014) compare the cyclicality properties of risk-based capital regulation with those of a non-risk-based required ratio of equity to total assets, which Basel regulators recently raised for larger banks. They find that the former tends to be more binding during economic upturns but to be less binding during downturns.

Do Tougher Capital Regulations Yield Better Outcomes? Empirical Evidence

Over roughly three decades of utilization of the Basel capital standards, initially primarily non-risk-based, leverage-ratio requirements relating assets to equity have

been complemented with risk-based requirements and supplemented with ratios including non-equity measures of regulatory capital. The most recent, manifestation of the Basel standards have re-emphasized equity-based requirements focused on total asset leverage. Except for one brief period prior to the recent crisis period, the general consequence has been considerable toughening of regulatory capital requirements.

Behr et al. (2010) examine 2006 data from more than 400 banks in 61 nations and find evidence of reduced risk taking—measured as the ratio of non-performing loans to total loans—in response to capital regulation in banking markets with low degrees of concentration. They find no such evidence in more concentrated banking markets. The authors' rationalization for this result is that in more concentrated markets, banks respond to tougher capital regulation by engaging in greater risk taking aimed at a long-run profitability objective incorporating a higher franchise value.

Jokipii and Milne (2011) examine data for most U.S. commercial banks between 1986 and 2008 and find contrasting relationships between bank capital adjustments and risk depending on banks' levels of capitalization. For banks with relatively high capital buffers, they find, risk and capital adjustments are positively related. In contrast, banks with relatively low levels of capital exhibit a negative relationship between the two variables.

Based on data from more than 3500 banks from 14 advanced nations over the 1998-period, Delis et al. (2012) find evidence of considerable heterogeneity of risk responses—changes in ratios of non-performing loans to total loans, *Z* scores, and volatility of return on assets—to capital regulation. Consistent with the basic results of Behr et al. (2010), Delis et al. find that banks possessing market power tend to respond to capital regulation by taking on more rather than less asset risk. They also find that more efficient banks are more likely to react by reducing their risk exposures.

The latter two studies' conclusions accord with the earlier, portfolio-based theoretical literature discussed earlier in the chapter. Their conclusions also are consistent with more recent analytical work discussed above that explicitly models differing choices across heterogeneous banks. Perhaps the contrasting influences of Basel capital standards on different groups of banks helps to explain the findings of Demirgüc-Kunt et al. (2013). These authors generally find little evidence of a relationship between the degree of bank capitalization and quarterly bank stock returns for more than 300 banks between 2005 and 2009. A specific exception was the interval of the global financial crisis, during which the relationship was strongest for the non-risk-based equity-assets ratio and for larger banks.

Does compliance with the broader range of Basel standards matter for bank risk? Demirgüc-Kunt and Detragiache (2011) explore this fundamental issue. They investigate whether basic compliance with per-Basel core principles was associated with *Z*-score measures of bank risk for more than 3000 banks in 86 countries between 1999 and 2006. They find no evidence of any relationship between Basel compliance and bank risk.

Ultimately, there is little evidence of unambiguously broad-based and unidirectional effects across a number of years of implementation of Basel capital standards. Why, one might ask, do national regulators persist in toughening and enlarging a complicated web of capital requirements that so far have generated such mixed or ambiguous outcomes? Posner (2015) concludes that the entire enterprise of capital regulation has involved "norming" on the part of bank regulators—the application of incremental changes in capital requirements in an effort to influence a handful of outlier institutions. He for replacing this norming approach to capital requirements should be replaced with a full-fledged cost-benefit-based approach to regulation intended to address risks across the banking industry as a whole. In actuality, however, the scope and complexity of the Basel capital-standards apparatus have continued to expand.

References

Acharya, Viral. 2009. A theory of systemic risk and design of prudential bank regulation. *Journal of Financial Stability* 5: 224–255.

Admati, Anat, and Martin Hellwig. 2014. *The Bankers' New Clothes: What's Wrong with Banking and What to Do About It*. Princeton, NJ: Princeton University Press.

Admati, Anat, Peter DeMarzo, Martin Hellwig, and Paul Pfleiderer. 2013. Fallacies, irrelevant facts, and mythgs in the discussion of capital regulation: Why bank equity is *not* expensive. Stanford University Rock Center for Corporate Governance, Working Paper Series No. 161, October 22.

Agénor, Pierre-Richard, and Luiz Perera da Silva. 2012. Cyclical effects of bank capital requirements with imperfect credit markets. *Journal of Financial Stability* 8: 43–56.

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 26–17.

Aiyar, Shekhar, Charles Calomiris, John Hooley, Yevgeniya Korniyenko, and Tomasz Wieladek. 2014a. The international transmission of bank capital requirements: Evidence from the UK. *Journal of Financial Economics* 113: 368–382.

Aiyar, Shekhar, Charles Calomiris, and Tomasz Wieladek. 2014b. Identifying credit substitution channels. *Economic Policy*, January, 45–77.

Aiyar, Shekhar, Charles Calomiris, and Tomasz Wieldek. 2015. How to strengthen the regulation of bank capital: Theory, evidence, and a proposal. *Journal of Applied Corporate Finance* 27: 27–36.

Akhigbe, Aigbe, Jeff Madura, and Marek Marciniak. 2012. Bank capital and exposure to the financial crisis. *Journal of Economics and Business* 64: 377–392.

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, Elena Carletti, and Robert Marquez. 2011. Credit market competition and capital regulation. *Review of Financial Studies* 24: 983–1018.

Almazan, Andres. 2002. A model of competition in banking: Bank capital versus expertise. *Journal of Financial Intermediation* 11: 87–121.

Anderson, Hanrik. 2011. Procyclical implications of Basel II: Can the cyclicality of capital requirements be contained? *Journal of Financial Stability* 7: 138–154.

Angkinand, Apanard. 2009. Banking regulation and the output cost of banking crises. *Journal of International Financial Markets, Institutions, and Money* 19: 240–257.

Ashcraft, Adam. 2001. Do tougher bank capital requirements matter? New evidence from the eighties. Manuscript, Federal Reserve Bank of New York, July 23.

Athanasoglou, Panayiotis, Ioannis Daniilidis, and Manthos Delis. 2014. Bank procyclicatlity and output: Issues and policies. *Journal of Economics and Business* 72: 58–83.

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.

Barakova, Irina, and Ajay Palvia. 2014. Do banks' internal Basle risk estimates reflect risk? *Journal of Financial Stability* 13: 167–179.

Barnea, Emanuel, and Moshe Kim. 2014. Dynamics of banks' capital accumulation. *Journal of Money, Credit, and Banking* 46: 779–816.

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.

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.

Behr, Patrick, Reinhard Schmidt, and Ru Xie. 2010. Market structure, capital regulation, and bank risk taking, 2010. *Journal of Financial Services Research* 37: 131–158.

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, 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.

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, Robert DeYoung, Mark Flannery, David Lee, and Özde Öztekin. 2008. How do large banking organizations manage their capital ratios? *Journal of Financial Services Research* 34: 123–149.

Besanko, David, and George Kanatas. 1996. The regulation of bank capital: Do capital standards promote bank safety? *Journal of Financial Intermediation* 5: 160–183.

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, and George Kaufman. 2003. Bank procyclicality, credit crunches, and asymmetric monetary policy effects: A unifying model. *Journal of Applied Finance* 2003 (Fall/Winter): 23–31.

Blum, Jürg. 1999. Do capital adequacy requirements reduce risks in banking? *Journal of Banking and Finance* 23: 755–771.

Blum, Jürg. 2008. Why 'Basel II' may need a leverage ratio restriction. *Journal of Banking and Finance* 32: 1699–1707.

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.

Borio, Claudio. 2003. Towards a macroprudential framework for financial supervision and regulation? BIS Working Paper No. 128, Bank for International Settlements, February.

Bouvatier, Vincent, and Laetitia Lepetit. 2008. Banks' procyclical behavior: Does provisioning matter? *Journal of International Financial Markets, Institutions, and Money* 18: 513–526.

Bradley, Michael, Carol Wambeke, and David Whidbee. 1991. Risk weights, risk-based capital, and deposit insurance. *Journal of Banking and Finance* 15: 875–893.

Brei, Michael, and Leonardo Gambacorta. 2014. The leverage ratio over the cycle. Bank for International Settlements Working Paper No. 471, November.

Brinkmann, Emile, and Paul Horvitz. 1995. Risk-based capital standards and the credit crunch. *Journal of Money, Credit, and Banking* 27: 848–863.

Brewer, Elijah III, George Kaufman, and Larry Wall. 2008. Bank capital ratios across countries: Why do they vary? *Journal of Financial Services Research* 34: 177–201.

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. 2013. Reforming banks without destroying their productivity and value. *Journal of Applied Corporate Finance* 25: 14–20.

Calomiris, Charles, and Charles Kahn. 1991. The role of demandable debt in structuring optimal banking arrangements. *American Economic Review* 81: 497–513.

Campbell, Tim, Yuk-Shee Chan, and Anthony Marino. 1992. An incentive-based theory of bank regulation. *Journal of Financial Intermediation* 2: 255–276.

Carey, Mark. 2002. A guide to choosing absolute bank capital requirements. *Journal of Banking and Finance* 26: 929–951.

Carlson, Mark, Hui Shan, and Missaka Warusawitharana. 2013. Capital ratios and bank lending: A matched bank approach. *Journal of Financial Intermediation* 22: 663–687.

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.

Chen, Andrew, Ju Nengjiu, Sumon Mazumdar, and Avinash Verma. 2006. Correlated default risks and bank regulations. *Journal of Money, Credit, and Banking* 38: 375–398.

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.

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.

Dagher, Jihad, Giovanni Dell'Ariccia, Luc Laeven, Lev Ratnovski, and Hui Tong. 2016. Benefits and costs of bank capital. IMF Staff Discussion Note No. SDN/16/04, March.

De Jonghe, Olivier, and Özde Öztekin. 2015. Bank capital management: International evidence. *Journal of Financial Intermediation* 24: 154–177.

Delis, Manthos, Kien Tran, and Efthymios Tsionas. 2012. Quantifying and explaining parameter heterogeneity in the capital regulation-bank risk nexus. *Journal of Financial Stability* 8: 57–68.

Demirgüc-Kunt, Asli, and Enrica Detragiache. 2011. Basel core principles and bank soundness: Does compliance matter? *Journal of Financial Stability* 7: 179–190.

Demirgüc-Kunt, Asli, Enrica Detragiache, and Ouarda Merrouche. 2013. Bank capital: Lessons from the financial crisis. *Journal of Money, Credit, and Banking* 45: 1147–1164.

Dermine, Jean. 2015. Basel III leverage ratio requirement and the probability of bank runs. *Journal of Banking and Finance* 53: 266–277.

Dia, Enzo, and David VanHoose. 2017. Fixed costs and capital regulation: Impacts on endogenous entry and aggregate loan quality. Unpublished Manuscript, Università degli Studi di Milano Bicocca and Baylor University.

Diamond, Douglas. 1984. Financial intermediation and delegated monitoring. *Review of Economic Studies* 51: 393–414.

Diamond, Douglas. 1991. Monitoring and regulation: The choice between bank loans and directly placed debt. *Journal of Political Economy* 99: 689–721.

Diamond, Douglas, and Phillip Dybvig. 1983. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy* 91: 401–419.

Diamond, Douglas, and Raghuram Rajan. 2000. A theory of bank capital. *Journal of Finance* 55: 2431–2465.

Diamond, Douglas, and Raghuram Rajan. 2001. Liquidity risk, liquidity creation, and financial fragility: A theory of banking. *Journal of Political Economy* 109: 287–327.

Dietrich, J. Kimball, and Christopher James. 1983. Regulation and the determination of bank capital changes: A note. *Journal of Finance* 38: 1651–1658.

Dowd, Kevin. 2000. Bank capital adequacy versus deposit insurance. *Journal of Financial Services Research* 17: 7–15.

Drehmann, Mathias, Claudio Borio, Leonardo Gambacorta, Gabriel Jiménez. 2010. Countercyclical capital buffers: Exploring options. Bank for International Settlements Working Paper No. 317, July.

Drumond, Ines. 2009. Bank capital requirements, business cycle fluctuations, and the Basel accords: A synthesis. *Journal of Economic Surveys* 23: 798–830.

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.

Fegatelli, Paolo. 2010. The misconception of the option value of deposit insurance and the efficacy of non-risk-based capital requirements in the literature on bank capital regulation. *Journal of Financial Stability* 6: 79–84.

Flannery, Mark. 1989. Capital regulation and insured banks' choice of individual loan default risks. *Journal of Monetary Economics* 24: 235–258.

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.

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* 3: 23–33.

Furlong, Frederick, and Michael Keeley. 1989. Capital regulation and bank risk-taking: A note. *Journal of Banking and Finance* 13: 883–891.

Gale, Douglas. 2010. Capital regulation and risk sharing. *International Journal of Central Banking* 6: 187–204.

Gennotte, Gerard, and David Pyle. 1991. Capital controls and bank risk. *Journal of Banking and Finance* 15: 805–824.

Gjerde, Øystein, and Kristian Semmen. 1995. Risk-based capital requirements and bank portfolio risk. *Journal of Banking and Finance* 19: 1159–1173.

Goodhart, Charles, Boris Hofmann, and Miguel Segoviano. 2004. Bank regulation and macroeconomic fluctuations. *Oxford Review of Economic Policy* 20: 591–615.

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.

Guidara, Alaa, Van Son Lai, Issouf Soumaré, and Fulbert Txchana Tchana. 2013. Banks' capital buffer, risk, and performance in the Canadian system: Impact of business cycle and regulatory changes. *Journal of Banking and Finance* 37: 3373–3387.

Hakenes, Hendrik, and Isabel Schnabel. 2011. Bank size and risk-taking under Basel II. *Journal of Banking and Finance* 35: 1436–1449.

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.

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.

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.

Holmström, Bengt, and Jean Tirole. 1997. Financial intermediation, loanable funds, and the real sector. *Quarterly Journal of Economics* 112: 663–691.

Homölle, Susanne. 2004. Bank capital regulation, asset risk, and subordinated uninsured debt. *Journal of Economics and Business* 56: 443–468.

Hotelling, Harold. 1929. Stability in competition. *Economic Journal* 39: 41–57.

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.

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.

Jeitschko, Thomas, and Shin Dong Jeung. 2005. Incentives for risk-taking in banking: A unified approach. *Journal of Banking and Finance* 29: 759–777.

Jokipii, Terhi, and Alistair Milne. 2011. Bank capital buffer and risk adjustment decisions. *Journal of Financial Stability* 7: 165–178.

Jokivuolle, Esa, Ilkka Kiema, and Timo Vesala. 2014. Why do we need countercyclical capital requirements? *Journal of Financial Research* 46: 55–76.

Jokivuolle, Esa, Jarmo Pesola, and Matti Viren. 2015. Why is credit-to-GDP a good measure for setting countercyclical capital buffers? *Journal of Financial Stability* 18: 117–128.

Kahane, Yehuda. 1977. Capital adequacy and the regulation of financial intermediaries. *Journal of Banking and Finance* 1: 207–218.

Kashyap, Anil, and Jeremy Stein. 2004. Cyclical implications of the Basel II capital standards. *Federal Reserve Bank of Chicago Economic Perspectives*, 18–31.

Keeley, Michael. 1988. Bank capital regulation in the 1980s: Effective or ineffective? *Federal Reserve Bank of San Francisco Economic Review*, Winter.

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.

Koehn, Michael, and Anthony Santomero. 1980. Regulation of bank capital and portfolio risk. *Journal of Finance* 35: 1235–1244.

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.

Kopecky, Kenneth, and David VanHoose. 2006. Capital regulation, heterogeneous monitoring costs, and aggregate loan quality. *Journal of Banking and Finance* 30: 2235–2255.

Kopecky, Kenneth, and David VanHoose. 2012. Can capital requirements induce private monitoring that is socially optimal? *Journal of Financial Stability* 8: 252–262.

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.

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.

Leung, W.S., N. Taylor, and K.P. Evans. 2015. The determinants of bank risks: Evidence from the recent financial crisis. *Journal of International Financial Markets, Institutions, and Money* 34: 277–293.

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.

Marini, François. 2003. Bank insolvency, deposit insurance, and capital adequacy. *Journal of Financial Services Research* 24: 67–78.

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.

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.

Memmel, Christoph, and Peter Raupach. 2010. How do banks adjust their capital ratios? *Journal of Financial Intermediation* 19: 509–528.

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.

Morrison, Alan, and Lucy White. 2005. Crises and capital requirements in banking. *American Economic Review* 95: 1548–1572.

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.

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.

Perotti, Enrico, Lev Ratnovski, and Razvan Vlahu. 2011. Capital regulation and tail risk. *International Journal of Central Banking*, 123–163.

Posner, Eric. 2015. How do bank regulators determine capital-adequacy requirements? *University of Chicago Law Review* 82: 1853–1895.

Repullo, Rafael. 2004. Capital requirements, market power, and risk-taking in banking. *Journal of Financial Intermediation* 13: 156–182.

Repullo, Rafael. 2013. Cyclical adjustment of capital requirements: A simple framework. *Journal of Financial Intermediation* 22: 608–626.

Repullo, Rafael, and Javier Suarez. 2004. Loan pricing under Basel capital requirements. *Journal of Financial Intermediation* 13: 496–521.

Rime, Bertrand. 2001. Capital requirements and bank behavior: Empirical evidence for Switzerland. *Journal of Banking and Finance* 25: 789–805.

Rochet, Jean-Charles. 1992. Capital requirements and the behavior of commercial banks. *European Economic Review* 36: 1137–1178.

Santos, João. 1999. Bank capital and equity investment regulations. *Journal of Banking and Finance* 23: 1095–1120.

Santos, João. 2001. Bank capital regulation in contemporary banking theory: A review of the literature. *Financial Markets, Institutions, and Instruments* 10: 41–84.

Schaeck, Klaus, and Martin Čihák. 2012. Banking competition and capital ratios. *European Financial Management* 18: 836–866.

Schliephake, Eva, and Roland Kirstein. 2013. Strategic effects of regulatory capital requirements in imperfect banking competition. *Journal of Money, Credit, and Banking* 45: 675–700.

Shim, Jeungbo. 2013. Bank capital buffer and portfolio risk: The influence of business cycle and revenue diversification. *Journal of Banking and Finance* 37: 761–772.

Shrieves, Ronald, and Drew Dahl. 1992. The relationship between risk and capital at commercial banks. *Journal of Banking and Finance* 16: 439–457.

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.

Stolz, Stéphanie. 2007. *Bank Capital and Risk-Taking: The Impact of Capital Regulation, Charter Value, and the Business Cycle*. Berlin: Springer.

Thakor, Anjan. 1996. Capital requirements, monetary policy, and aggregate bank lending: Theory and empirical evidence. *Journal of Finance* 51: 279–324.

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.

VanHoose, David. 2008. Bank capital regulation, economic stability, and monetary policy: What does the academic literature tell us? *Atlantic Economic Journal* 36: 1–14.

Wagster, John. 1999. The Basel accord of 1988 and the international credit crunch of 1989–1992. *Journal of Financial Services Research* 15: 123–143.

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.

Wang, Lang. 2005. Bank capital requirements and the effectiveness of monetary policy. Manuscript, Peking University.

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). The primary objective between during 1980s through the early 2000s was placed on prevention, culminating in the Basel I and II frameworks for international bank regulation and supervision. In spite of the spectacular failure of these frameworks to satisfy the prevention objective, the Basel III standards adopted in recent years double down on toughened requirements aimed at this same goal.

The framers of Basel II suggested an analogy in which the framework is an elevated foundation resting on three "pillars," which Basel III has retained 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% of the description of the Basel II framework (Bank for International Settlements 2006) covered capital requirements, and the bulk of Basel III changes involve capital and liquidity regulation. Hence, the brightest spotlight has shone on capital adequacy regulation.

Most policymakers and researchers agree that the Basel regulatory and supervisory framework 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 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 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 and supervisory discipline applied by

© Springer-Verlag GmbH Germany 2017

D. VanHoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-662-54326-9_8

government regulators. The chapter concludes by evaluating the market discipline pillar in relation to the Basel capital-standards and supervisory-process pillars.

The Basel Framework's Market Discipline Pillar

The basic structure of Basel of the three Basel pillars—developed for Basel II but retained under Basel III—is laid out in the Bank for International Settlements' "International Convergence of Capital Measurement and Capital Standards" (2006). Roughly 150 pages discuss 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 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, 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 failure 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 the Basel framework is the disclosure of information, particularly in relation to banks' capital positions. National supervisors are responsible for ensuring that banks establish policies for information disclosure either semi-annually or quarterly.

Left unaddressed is whether the Basel framework'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 the Basel system likely to promote "discipline" by the market? And how much are the guidelines likely to promote a safer and sounder banking system?

Lane (1993) describes market discipline as the provision of signals, typically in the form of cutbacks in sources of funds use to finance banks' asset portfolios, which provide incentives to engage in solvency-promoting actions. Banks obtain funds by issuing deposits, other debt subordinate to deposits in the event of failure, and equity-ownership shares 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. 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. Hasman et al. (2013) reach an analogous conclusion about the potential benefits of depositor discipline based on analysis of a theoretical model of deposit contracting between potentially undercapitalized banks and customers weighing desired returns versus the potential for runs.

Consistent with these view, a study of 1994–2013 data from U.S. bank holding companies by Alanis et al. (2015) concludes that a one-standard-deviation increase in uninsured deposits pushes down by 46–64 basis points the required bond yield on banks' subordinated debts. Alanis et al. argue that this finding indicates a market perception that uninsured depositors discipline banks. Barajas and Catalán (2015) find evidence that large Argentine pension-fund depositors disciplined banks between 1998 and 2001 but only in the absence of conflicts of interest involving bank-owned pension fund management companies. Berger and Turk-Ariss (2015) utilize financial-statement data for more than 2000 U.S. and European banks over the 1997–2009 interval to examine the relationship between deposit growth and bank risk-taking measured in terms of the equity-asset ratio and the ratio of nonperforming loans to total loans. Their results indicate the presence of depositor discipline, particularly for larger U.S. banks, prior to the global financial crisis but an absence of such discipline during the crisis period, except at smaller U.S. banks.

Because bank equity owners gain if the bank experiences gains from risk taking, they more conflicted motivations regarding the bank's exposure to risk. These agents consider expected profits and risk simultaneously and take into account debt holders' responses to bank risks. 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.

Two factors complicate the equity-ownership-based channel of market discipline. One of these is dispersion of ownership that reduces incentives for and capabilities of individual shareholders to discipline managers. Analysis of quarterly 1997–2005 data from 76 banks in 18 European nations by Auvray and Brossard (2012) offers evidence that an equity-ownership disciplining effect requires individual shareholders' ownership of sufficiently large blocks of shares, and Caner et al. (2012) utilize quarter data from Turkish banks over the 1997–2006 interval to determine that small banks with less ownership dispersion exercise greater market discipline over managers than is true at larger banks. In a study of 1995–2005 data from more than 300 banks in 47 nations, Forssbæk (2011) concludes that a greater degree of control on the part of equity owners strengthens the risk-reducing effects of other sources of market discipline on banks.

The other factor is the potential moral-hazard effect of deposit insurance on owners' preferences. In principle, this effect could induce bank equity owners to offer compensation to managers for taking on excessive risk—for recent perspectives, see Gregg et al. (2012), Thanassoulis (2014), and Thanassoulis and Tanaka (2015)—which has engendered a growing literature on the regulation of bank management compensation—see, for instance, Chaigneau (2013) and Hakenes

and Schnabel (2014). As discussed in VanHoose (2011a), however, the empirical evidence on the presence and magnitude of the effect of bank-management compensation on risk is mixed, and contrasting implications of compensation regulation for bank outcomes have not yet been fully assessed. Nevertheless, an application of a very basic theory of equilibrium choices across the banking industry's external margin in VanHoose (2011b) indicates likely reductions in bank efficiency and potential unintended decreases in loan quality.

Conditions for Market Signals to Effectively Discipline Banks

The above discussion suggests that suppliers of funds 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 monitoring activities by funds suppliers and perhaps even regulators feeds back to affect prices of banks' secondary debt instruments, "indirect" market influence may induce corrective action by banks.

Several conditions must be satisfied for depositors and debt and equity owners 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, they must have information about bank management, capital structure, and risk exposures.

Second, 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 managers to respond by reducing risk exposures or for equity owners to replace the current managers with new ones 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 et al. (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 speedy resolutions 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 perceive 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 debt and equity markets could have advantages over banks in less developed countries with relatively closed and inactive markets. 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 bank market discipline requires market participants to be informed about risk profiles. Immediate knowledge of risks is private information to banks' managers, however, 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 managers could be costly and thus are willing to accept some discretion for managers to withhold a "threshold" quantity of information.

In addition, Boot and Thakor (2001) provide an analysis suggesting that the quantity and type of private managerial information to disclose depends on whether it complements or substitutes for public information. 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. 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 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 disclosure faces a trade-off: Weaker disclosure standards encourage socially costly expropriations by firm insiders, but stricter standards reduce the returns of the insiders and thereby discourage their investment. Thus, mandatory 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 transparency tends to boost the likelihood of bank runs, unless 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. They show that Calomiris and Kahn (1991) describe how sophisticated sellers of wholesale funds can add value by monitoring banks and imposing market discipline, 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.

The theoretical framework provided by Vauhkonen (2012) builds on Hellmann et al.'s (2000) and Repullo's (2004) model of bank capital regulation and is set within a spatial-competition model with bank insiders and outside investors. As noted in Chap. 7, this class of models tends to favor tougher regulatory requirements, and that is likewise the case here. In the model, banks find any voluntary disclosure prohibitively costly, but imposing mandatory disclosure unambiguously makes the banking system safer.

Van Tassel (2011) analyzes a theoretical model within which banks can disclose credit-market information in advance of competition for borrowers. To the extent it discloses such information, the bank gives up an information advantage in credit markets, but it gains an advantage because of the resulting drop in the rate at which it funds its lending. Thus, like Östberg's model, Van Tassel's implies a trade-off that yields mixed incentives for information disclosure. In Van Tassel's framework, the disclosure incentive depends negatively on the bank's capital ratio. There is a positive network effect, however, in that an individual bank has a greater incentive to disclose information as more competing banks do so.

Moreno and Takalo (2016) consider a theoretical setting in which competitive banks fund acquisitions of illiquid assets via short-term borrowings that must periodically be rolled over. Their model also yields a trade-off—in this case, between greater transparency contributing to more efficient liquidation but at the same time magnifying rollover risks. Moreno and Takalo consequently conclude that the optimal degree of bank disclosure implies neither full transparency nor complete opacity.

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 if investors have full information and market signals are clear to all participants, markets will fail to discipline banks if government policies insulate managers from market signals. 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.

Saunders observes that explicit U.S. legal changes in the early 1990s expressly exempted subordinated debts from too-big-to-fail guarantees. Thus, suppliers of funds 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 debts 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 (2008) 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.

Nevertheless, although Belkhir's (2013) analysis of quarterly 1995–2009 U.S. bank call report data indicates that the relative magnitude of banks' issuances of subordinated debts does influence risk management through interest-rate derivative securities in a manner consistent a with market-disciplining effect, such an effect is missing at too-big-to-fail banks. This finding is consistent with that of Balasubramnian and Cyree (2011), who estimate size- and too-big-to-fail-based discounts on bank yield spreads for more than a decade following the 1998 bailout of Long-Term Capital Management. In a study of banking data from 90 nations from 1991 to 2011, Bertay et al. (2013) determine that large banks' funding costs are subject to market-discipline effects, in that those costs become more sensitive to measures of the banks' risk positions as the banks increase in size—which they interpret as an indication that banks can become too large to justify saving. At the same time, however, Bertay et al. find that the funding costs diminish with size, so market discipline does not prevent bank expansions. Beyhaghi et al. (2014) examine data from 6 large and 15 small Canadian banks between 1990 and 2010 and find that although market discipline operates through subordinated debts, large banks have funding advantages. They pay 80 basis points less per dollar of deposits and 70 basis points less per dollar of subordinated debt.

Separately, Balasubramnian and Cyree (2014) conclude that discounts on large banks' yield-spread discounts have decreased substantially during the current decade. They suggest that these reductions may have resulted from enforcement of various provisions of the Dodd-Frank Act of 2010.

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 mandating issuance of bank subordinated debt, which Zhang et al. (2014) show to reflect sensitivity to banks' risks, 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), Bliss (2001), and Evanoff et al. (2011)]. 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 1–5 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 offered by Lang and Robertson (2002) that such an approach would replace regulators' subjective judgment with market signals without boosting the risk sensitivity of bank costs. Calomiris and Powell (2000) argue that this approach would eliminate "plausible deniability" for regulatory inaction and give 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) find that subordinated debt spreads are most consistent 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) 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%.

Herring (2004) notes that integrating market discipline into bank regulatory requirements has the potential to raise 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? Prompt supervisory actions make failure more likely. Yet a more gradual approach to corrective action could give the managers more time to weaken an institution further and increase closing costs; if so, 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.

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, $\widehat{E}^i$, given by $\widehat{E}^i = P\left(E^i\right)\Pi - C^i\left(L^i, S^i, D^i, N^i, E^i\right)$, where E^i denotes beginning-of-period equity capital, which is assumed to be unregulated, and where $\Pi = R_L^i\left(L^i\right) L^i + R_S S^i - R_D^i\left(D^i\right)D^i - R_N^i\left(N^i, E^i\right)N^i$ defines gross interest earnings net of gross interest expenses, with $R_L^i = 1 + r_L^i$, $R_S = 1 + r_S$, $R_D^i = 1 + r_D^i$, and $R_N^i = 1 + r_N^i$. $P(E)$ 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 nonnegative 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:

$$
\begin{aligned}
&P(E^i)\left[1-(\varepsilon^i)^{-1}\right]R_L^i - C_L^i = P(E^i)R_S - C_S^i = P(E^i)\left[1+(\mu^i)^{-1}\right]R_N^i + C_N^i \\
&= (1-q)^{-1}\left\{\left[1+(\eta^i)^{-1}\right]P(E^i)R_D^i + C_D^i\right\} \\
&= P'(E^i)\Pi + P(E^i)R_N^i\left(1-\xi^i\frac{N^i}{E^i}\right) + C_E^i,
\end{aligned}
$$

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 $\xi^i \equiv \frac{\left(\partial R_N^i/R_N^i\right)}{\left(\partial E^i/E^i\right)}$ is the absolute elasticity of the rate of return on non-deposit liabilities with respect to beginning-of-period equity. 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 equity.

Note that because $\partial R_N^i/\partial E^i < 0$, an increase in beginning-of-period equity 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 reduced marginal cost of beginning-of-period equity. 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 more 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 lower 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 incumbents with interdependent investment strategies to offer less informative disclosure 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 under specialized and universal banking arrangements. With 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 literature has had relatively little say about the effects of variations in market power on bank market discipline. This issue can be examined within the Landskroner-Paroush framework. Note that the spread between the rates of return on non-deposit liabilities—such as subordinated debts—and on deposits in terms of the exogenously determined market security rate can be expressed as

$$\frac{R_N^i - R_D^i}{R_D^i} = \frac{\left[P\left(E^i\right)R_S - C_S^i - C_N^i\right]\left[1 + \left(\eta^i\right)^{-1}\right]}{\left\{(1-q)\left[P\left(E^i\right)R_S - C_S^i\right] - C_D^i\right\}\left[1 + \left(\mu^i\right)^{-1}\right]} - 1.$$

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 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:

$$P(E^i)\left[1 + (\mu^i)^{-1}\right]R_N^i + C_N^i = P(E^i)R_N^i\left(1 - \xi^i\frac{N^i}{E^i}\right) + C_E^i.$$

This in turn implies that at an optimum, the degree of market discipline is $\xi^{i*} = 1 - (\mu^i)^{-1}\frac{E^i}{N^i}$. 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 rate-of-return variations 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 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 $\mu^i \to \infty$, ξ^{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.

Andrievskaya and Semenova (2016) argue that how much information banks disclose can feed back to influence concentration and market power. On the one hand, increased transparency can reveal some strategic information that can promote competition. In addition, binding minimal disclosure requirements can reduce the market-power gains that otherwise might have been forthcoming if banks could withhold information strategically, plus better-informed potential entrants might be more willing to enter banking markets, thereby contributing to greater competition. On the other hand, greater information disclosure might enable facilitation of coordinated actions by banks that could contribute, for instance, to a tighter oligopoly. Furthermore, rules requiring binding minimal required transparency could boost banks' fixed compliance costs and thereby raise minimum efficient scale, which could boost concentration and market power.

Based on 1998, 2001, 2005, and 2010 data for dozens of nations, Andrievskaya and Semenova find that for a given level of credit risk, the level of concentration of banking markets is negatively related to bank transparency. With sufficiently high credit quality, however, this relationship tends to weaken, so the relationship is stronger for less advanced and emerging nations. Their two measures of the degree of competition—the interest margin and the Lerner index—yield mixed effects for greater disclosure.

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). Let's consider the evidence.

Information Content of Market Prices and Bond Yield Spreads

A key issue is whether funds suppliers can perceive changes in banks' risk profiles. Flannery (1998) provides a review of the evidence through the mid-1990s and concludes that there is a role for market discipline in supplementing supervision. Flannery and Sorescu (1996) find that investors can distinguish among risks taken by major U.S. banks, especially during periods when subordinated debt is not covered by government guarantees. They find that spreads between yields on subordinated debtz and Treasury securities with the same maturities can 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 firms. Stocks of smaller bank holding companies trade much less frequently but have spreads similar to those of matched nonfinancial firms. Bank holding companies' stocks appear to be no more informationally opaque than other stocks, although 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.

Greater opaqueness of banking could help to explain Alsakka et al.'s (2014) finding, derived from analysis of 2003–2013 data, that credit-rating agencies initially were slower during the financial crisis period to downgrade ratings of banks' debts than those of sovereign nations but later lower bank ratings more sharply. Bank opaqueness certainly helps to explain Anolli et al.'s (2014) finding, based on a study of 36,232 forecasts by bank financial analysts, that bank-specific risks contributed negatively to the ability of the financial analysts to gauge risks between 2003 and 2009.

Morgan and Stiroh (2001) examine spreads on nearly 500 new bonds issued by U.S. banks and 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 asset portfolios, although banks are most likely to enter primary debt markets when perceptions of bank risks are more favorable. These results apply to the 1993–1998 period of good times for U.S. banks (1993–1998), so Morgan and Stiroh conclude that debt markets provide clear signals of asset risk differentials 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 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 yields 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 et al. (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, Iannotta's (2006) study of bond spreads for more than 100 - European banks between 1999 and 2007 suggests that investors are able to discern hidden information about banks' performances and prospects.

Nevertheless, work by Berger and Davies (1998) and DeYoung and Roland (2001) find that 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 and Roland 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 6 months) and that subordinated debenture prices do not immediately reflect this information. DeYoung and Roland argue that when information obtained by supervisory examinations is released, market prices reflect not only the information itself but also likely regulatory actions.

Should regulators release all information obtained from examinations? Prescott (2008) examines a model involving interactions among a bank, a 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.

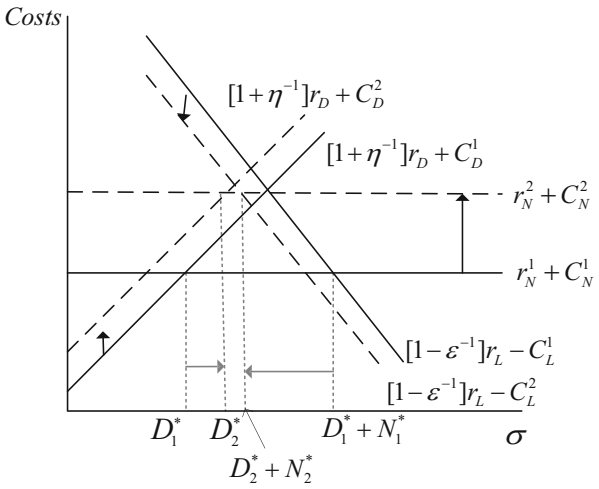

Fig. 8.1 Combined effects of increased bank risk with regulatory and market discipline reactions

Regulatory Crowding Out of Market Discipline?

Billett et al. (1995) provide an analysis suggesting that regulatory actions in response to new information about risks give imperfectly competitive banks incentives to reshuffle liabilities to protect themselves from market discipline. Figure 8.1 illustrates the argument of Billett 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 deposits and nondeposit liabilities and the total of the two, which are assumed to constitute the bank's total liabilities.

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). The market rate of return on non-deposit liabilities also rises 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 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. (1998) examine data on 116 downgrades in credit ratings of US. bank holding companies between 1990 and 1995. They find that, consistent

with 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.

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. 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 result.

On the basis of an analysis of 79 different banking crises involving 66 nations, Cubillas et al. (2012) find that expansions of deposit insurance and other accommodative policies undertaken during such events tend to be associated with declines in market discipline. In a study of an expansion of deposit insurance to Russian households in 2004 and a later minor banking crisis, Karas et al. (2013) likewise find a reduction in t market discipline.

Cihak et al. (2013) study, within the context of data for 143 nations, regulatory and supervisory responses to the global financial crisis. They find that capital regulation toughened, structures for bank governance were revamped, and rules governing crisis resolution were strengthened. Consistent with the regulatory crowding-out perspective, however, Cihak et al. conclude that private market discipline diminished.

Li and Song (2013) utilize data from 277 banks in 55 nations to consider the effects of broadened powers for governmental regulators to monitor and discipline banks. They find evidence that the expansion of such powers reduces the independence of banks' governing boards. The result, they conclude, is a decrease in the efficiency of market discipline.

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 et al. (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.

Only limited evidence exists regarding 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% 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.

As discussed later in this chapter, theoretical arguments regarding the bank- or societal-level pros and cons of bank information disclosure are mixed. Nevertheless, requiring banks to formulate processes for disclosure of information is at the heart of the Basel market discipline pillar. 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 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 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 capital positions, particularly in countries with more competitive banking systems.

Evidence on Bank Information Disclosure

Jordan et al. (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.

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 a 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. In addition, Demirgüc-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 Basel information-disclosure requirements were more financially sound.

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. Much less evidence exists about uninsured depositors' use of and responses to information, but what we have indicates that depositors, like investors, use and react to available information.

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 the Basel Framework

A key problem in evaluating the effects of the informational requirements of the Basel framework 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 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 the Basel Framework

The Basel 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 et al. (2003) observe, the Basel market discipline pillar 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. It 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 framers contemplated the *sufficient* conditions to enable markets to discipline via clear market signals and effective monitoring.

The Basel 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 the Basel system could result in dramatic improvements.

Yet the Basel framework 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 the Basel Market Discipline Pillar

Very little research has analyzed the interactions among the three Basel pillars, although Llewellyn and Mayes (2003) offer common-sense conditions for market discipline and regulatory prompt corrective action to be complementary. Decamps et al. (2004), use a continuous-time model to analyze implied market and book values of a bank with and without capital requirements, regulatory auditing, and market discipline. 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. If the monitoring cost is relatively "low," there is a range of outcomes in which the market value of using the bad technology exceeds that for the good technology, in which case a bank evades monitoring and increases its likelihood of failure. A sufficiently high monitoring cost induces bank shareholders voluntarily shut down the bank. Decamps et al. consider the 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 it will opt for the "bad technology." The regulator can 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-style three-pillar mix—provided that securities prices are not so volatile that they fail to yield information about actual cash flows.

Decamps et al.'s conclusions hinge on two key assumptions. One is that regulators *require* banks to hold subordinated debt, which reduces the bank's market value but allows for a lower capital requirement while inducing the bank to choose the good technology. They also assume that regulators are free from political interference and follow *rules* for bank closures. Under these assumptions, the three pillars of capital regulation, market discipline, and a supervisory process can be reinforcing.

The Basel 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, the Basel system'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. Hence, the Basel framework pays insufficient attention to the useful regulatory role of market discipline and relies inappropriately on supervisory discretion—a key topic of the next chapter.

References

Alanis, Emmanuel, Hamid Beladi, and Margot Quijano. 2015. Uninsured deposits as a monitoring device: Their impact on bond yields of banks. *Journal of Banking and Finance* 52: 77–88.

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, 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.

Alsakka, Rasha, Owain ap Gwilym, and Tuyet Nhung Vu. 2014. The sovereign-bank rating channel and rating agencies' downgrades during the European debt crisis. *Journal of International Money and Finance* 49: 235–257.

Andrievskaya, Irina, and Maria Semenova. 2016. Does banking system transparency enhance bank competition? Cross-country evidence. *Journal of Financial Stability* 23: 33–50.

Anolli, Mario, Elena Beccalli, and Philip Molyneux. 2014. Bank earnings forecasts, risk, and the crisis. *Journal of International Financial Markets, Instruments, and Money* 29: 309–335.

Ashcraft, Adam. 2008. Does the market discipline banks? New evidence from the regulatory capital mix. *Journal of Financial Intermediation* 17: 543–561.

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.

Auvray, Tristan, and Olivier Brossard. 2012. Too dispersed to monitor? Ownership dispersion, monitoring, and the prediction of bank distress. *Journal of Money, Credit, and Banking* 44: 685–714.

Balasubramnian, Bhanu, and Ken Cyree. 2011. Market discipline of banks: Why are yield spreads on bank-issued subordinated notes and debentures not sensitive to bank risks? *Journal of Banking and Finance* 35: 21–35.

Balasubramnian, Bhanu, and Ken Cyree. 2014. Has market discipline on banks improved after the Dodd-Frank Act? *Journal of Banking and Finance* 41: 155–166.

Bank for International Settlements. 2006. International convergence of capital measurement and capital standards: A revised framework comprehensive version, June.

Barajas, Adolfo, and Mario Catalán. 2015. Market discipline and conflicts of interest between banks and pension funds. *Journal of Financial Intermediation* 24: 411–440.

Barth, James, Gerard Caprio Jr, and Ross Levine. 2004. Bank regulation and supervision: What works best? *Journal of Financial Intermediation* 13: 205–248.

Barth, James, Gerard Caprio Jr, and Ross Levine. 2006. *Rethinking Bank Regulation*. Cambridge, UK: Cambridge University Press.

Belkhir, Mohamed. 2013. Do subordinated debt holders discipline bank risk-taking? Evidence from risk management decisions. *Journal of Financial Stability* 9: 705–719.

Berger, Allen, and Sally Davies. 1998. The information content of bank examinations. *Journal of Financial Services Research* 14: 117–144.

Berger, Allen, and Rima Turk-Ariss. 2015. Do depositors discipline banks and did government actions during the recent crisis reduce this discipline? An international perspective. *Journal of Financial Services Research* 48: 103–126.

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.

Bertay, Ata Can, Asli Demirgüç-Kunt, and Harry Huizinga. 2013. Do we need big banks? Evidence on performance, strategy, and market discipline. *Journal of Financial Intermediation* 22: 532–558.

Beyhaghi, Mehdi, Chris D'Souza, and Gordon Roberts. 2014. Funding advantage and market discipline in the Canadian banking sector. *Journal of Banking and Finance* 48: 396–410.

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.

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.

Birchler, Urs, and Andréa Maechler. 2001. Do depositors discipline Swiss banks? Swiss National Bank Working Paper No. 1.06.

Bliss, Robert. 2001. Market discipline and subordinated debt: A review of some salient issues. *Federal Reserve Bank of Chicago Economic Perspectives*, First Quarter: 24–45.

Bliss, Robert. 2004. Market discipline: Players, processes, and purposes. In *Market Discipline Across Countries and Industries*, ed. Claudio Borio, William Hunter, George Kaufman, and Kostas Tsatsaronis, 37–53. Cambridge, MA: MIT Press.

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.

Blum, Jürg. 2002. Subordinated debt, market discipline, and banks' risk taking. *Journal of Banking and Finance* 26: 1427–1441.

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 2004.

Bond, Philip, Itay Goldstein, and Edward Simpson Prescott. 2008. Market-based corrective action. Manuscript, University of Pennsylvania and Federal Reserve Bank of Richmond, July.

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.

Bordo, Michael. 2003. Market discipline and financial crisis policy: A historical perspective. In *Market Discipline in Banking: Theory and Evidence*, ed. George Kaufman, 157–182. Amsterdam: Elsevier.

Calomiris, Charles. 1999. Building an incentive-compatible safety net. *Journal of Banking and Finance* 23: 1499–1519.

Calomiris, Charles. 2004. Reestablishing market discipline as a part of bank regulation. In *Market Discipline Across Countries and Industries*, ed. Claudio Borio, William Hunter, George Kaufman, and Kostas Tsatsaronis, 407–421. Cambridge, MA: MIT Press.

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.

Caner, Selçuk, Süheyla Özyıldırım, and A. Ece Ungan. 2012. How sensitive are bank managers to shareholder value? *Journal of Financial Services Research* 42: 187–205.

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.

Chaigneau, Pierre. 2013. Risk-shifting and the regulation of bank CEOs' compensation. *Journal of Financial Stability* 9: 776–789.

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.

Cihak, Martin, Asli Demirgüç-Kunt, Maria Soledad Martinez Peria, and Amin Mohseni-Cheraghlou. 2013. Bank regulation and supervision in the context of the global crisis. *Journal of Financial Stability* 9: 733–746.

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.

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.

Cubillas, Elena, Ana Rosa Fonseca, and Francisco González. 2012. Banking crises and market discipline: International evidence. *Journal of Banking and Finance* 36: 2285–2298.

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.

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.

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.

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.

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.

Evanoff, Douglas, and Larry Wall. 2000. Subordinated debt as bank capital: A proposal for regulatory reform. *Federal Reserve Bank of Chicago Economic Perspectives* 24: 40–53.

Evanoff, Douglas, and Larry Wall. 2001. Sub-debt yield spreads as bank risk measures. *Journal of Financial Services Research* 20: 121–145.

Evanoff, Douglas, Julapa Jagtiani, and Taisuke Nakata. 2011. Enhancing market discipline in banking: The role of subordinated debt in financial regulatory reform. *Journal of Economics and Business* 63: 1–22.

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. 2001. The faces of "market discipline". *Journal of Financial Services Research* 20: 107–119.

Flannery, Mark, and Sorin Sorescu. 1996. Evidence of bank market discipline in subordinated debenture yields: 1983–1991. *Journal of Finance* 51: 1347–1377.

Flannery, Mark, Simon Kwan, and M. Nimalendran. 2004. Market evidence on the opaqueness of banking firms' assets. *Journal of Financial Economics* 71: 419–460.

Forssbæk, Jens. 2011. Ownership structure, market discipline, and banks' risk-taking incentives under deposit insurance. *Journal of Banking and Finance* 35: 2666–2678.

Gilbert, R. Alton. 1990. Market discipline of bank risk: Theory and evidence. *Federal Reserve Bank of St. Louis Review*, January–February 3–18.

Goyal, Vidhan. 2005. Market discipline of bank risk: Evidence from subordinated debt contracts. *Journal of Financial Intermediation* 14: 318–350.

Gregg, Paul, Sarah Jewell, and Ian Tonks. 2012. Executive pay and performance: Did bankers' bonuses cause the crisis? *International Review of Finance* 12: 89–122.

Gunther, Jeffery, and Robert Moore. 2003. Loss underreporting and the auditing role of bank exams. *Journal of Financial Intermediation* 12: 153–177.

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.

Hakenes, Hendrik, and Isabel Schnabel. 2014. Bank bonuses and bailouts. *Journal of Money, Credit, and Banking* 46: 259–288.

Hamalainen, Paul, Maximilian Hall, and Barry Howcroft. 2003. Market discipline: A theoretical framework for regulatory policy development. In *Market Discipline in Banking: Theory and Evidence*, ed. George Kaufman, 57–95. Amsterdam: Elsevier.

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.

Hasman, Augusto, Margarita Samartin, and Jos Van Bommel. 2013. Financial contagion and depositor monitoring. *Journal of Banking and Finance* 37: 3076–3084.

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.

Herring, Richard. 2004. How can the invisible hand strengthen prudential regulation and how can prudential regulation strengthen the invisible hand? In *Market Discipline Across Countries and Industries*, ed. Claudio Borio, William Hunter, George Kaufman, and Kostas Tsatsaronis, 363–379. Cambridge, MA: MIT Press.

Huang, Rocco, and Lev Ratnovski. 2008. The dark side of bank wholesale funding. Federal Reserve Bank of Philadelphia Working Paper No. 09-3, November.

Iannotta, Giuliano. 2006. Testing for opaqueness in the European banking industry: Evidence from bond credit ratings. *Journal of Financial Services Research* 30: 287–309.

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.

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.

Karas, Alexei, William Pyle, and Koen Schoors. 2013. Deposit insurance, banking crises, and market discipline: Evidence from a natural experiment on deposit flows and rates. *Journal of Money, Credit, and Banking* 45: 179–200.

King, Thomas. 2008. Discipline and liquidity in the interbank market. *Journal of Money, Credit, and Banking* 40: 295–317.

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. Board of Governors of the Federal Reserve System Staff Study 172.

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, and Douglas Robertson. 2002. Analysis of proposals for a minimum subordinated debt requirement. *Journal of Economics and Business* 54: 115–136.

Li, Li, and Frank Song. 2013. Do bank regulations affect board independence? A cross-country analysis. *Journal of Banking and Finance* 37: 2714–2732.

Llewellyn, David, and David Mayes. 2003. The role of market discipline in handling problem banks. In *Market Discipline in Banking: Theory and Evidence*, ed. George Kaufman, 183–210. Amsterdam: Elsevier.

Maclachlan, Fiona. 2001. Market discipline in bank regulation: Panacea or paradox? *The Independent Review* 6: 227–234.

McDill, Kathleen, and Andrea Maechler. 2003. Do uninsured depositors vote with their feet? In *Market Discipline in Banking: Theory and Evidence*, ed. George Kaufman, 211–248. Amsterdam: Elsevier.

Moreno, Diego, and Tuomas Takalo. 2016. Optimal bank transparency. *Journal of Money, Credit, and Banking* 48: 203–231.

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.

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.

Östberg, Per. 2006. Disclosure, investment, and regulation. *Journal of Financial Intermediation* 15: 285–306.

Park, Sangkyun, and Stavros Peristiani. 1998. Market discipline by thrift depositors. *Journal of Money, Credit, and Banking* 30: 347–364.

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.

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.

Podpiera, Richard. 2006. Does compliance with Basel core principles bring any measurable benefits? *IMF Staff Papers* 53: 306–326.

Prescott, Edward Simpson. 2008. Should bank supervisors disclose information about their banks? *Federal Reserve Bank of Richmond Economic Quarterly* 94: 1–16.

Repullo, Rafael. 2004. Capital requirements, market power, and risk-taking in banking. *Journal of Financial Intermediation* 13: 156–182.

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.

Saunders, Anthony. 2001. Comments on Evanoff and Wall/Hancock and Kwast. *Journal of Financial Services Research* 20: 189–194.

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.

Tadesse, Solomon. 2005. Banking fragility and disclosure: International evidence. Manuscript, University of South Carolina.

Thanassoulis, John. 2014. Bank pay caps, bank risk, and macroprudential regulation. *Journal of Banking and Finance* 48: 139–151.

Thanassoulis, John, and Misa Tanaka. 2015. Bankers' pay and excessive risk. Bank of England Staff Working Paper No. 558, October.

Van Tassel, Eric. 2011. Information disclosure in credit markets when banks' costs are endogenous. *Journal of Banking and Finance* 35: 490–497.

VanHoose, David. 2011a. Regulation of bank management compensation. In *Financial Market Regulation: Legislation and Implications*, ed. John Tatom, 163–183. New York: Springer.

VanHoose, David. 2011b. Regulatory constraints on performance-based managerial compensation, bank monitoring, and aggregate loan quality. *Atlantic Economic Journal* 39: 315–328.

Vauhkonen, Jukka. 2012. The impact of pillar 3 disclosure requirements on bank safety. *Journal of Financial Services Research* 41: 37–49.

Verrecchia, Robert. 1983. Discretionary disclosure. *Journal of Accounting and Economics* 5: 179–194.

Zhang, Zhichao, Wei Song, Xin Sun, and Nan Shi. 2014. Subordinated debt as an instrument of market discipline: Risk sensitivity of sub-debt yield spreads in UK banking. *Journal of Economics and Business* 73: 1–21.

Chapter 9
Regulation and the Structure of the Banking Industry

To this point, discussion of sources of interplay between bank market structure and regulation has emphasized how market structure influences effects of regulatory policies. This chapter focuses on how regulation itself 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.

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 principles textbooks, such as Miller (2016) 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 people place on these wants and desires. Thus, the notion of "needs," the textbooks suggest, has no place in discussions of economic issues.

© Springer-Verlag GmbH Germany 2017

D. VanHoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-662-54326-9_9

Yet researchers often allege a "need" for bank regulation. Santos (2001, p. 46) 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. . ." 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 of a "need" for regulation in order for any regulation to exist helps to explain why some economists insist that any analysis of bank regulation makes sense only in a setting in which regulation 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 so heavily regulate banks, particularly in light of Barth et al.'s (2014b) conclusion that following substantial changes in bank regulation and supervision between 1993 and 2013, large banks remain at least as dominant, interdependent banking risks have worsened, and moral hazard problems have broadened?

Public Choice Motivations for Bank Regulation

In fact, industrial organization economists have long understood that there are motivations for regulating firms that have nothing to do with aiming to correct market failures. Certainly, the existence of third-party spillover costs or benefits, inefficiencies arising from market power, and welfare losses from informational asymmetries can motivate 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 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, insisting that analysis of bank regulation makes sense only in models in which regulation is "needed" are out of tune with the field of regulatory economics. Viscusi et al. (2005, pp. 376–378) review the public interest theory of regulation, which they refer to as "normative analysis as positive theory." They note that it hypothesizes that regulation "occurs when it should occur because the potential for a net social welfare gain generates public demand for regulation"—

that is, a 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." They also contend that even a common reformulation in which regulation initially is established to correct a market failure but then often is mismanaged by regulatory agencies 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) in the case in which banks face no fixed costs and constant marginal resource costs. In this situation, as discussed in Chap. 2, 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, and elected representatives and thereby also generates utility for the regulator. Thus, the regulator's indifference curves slope upward, as depicted in panel (a),

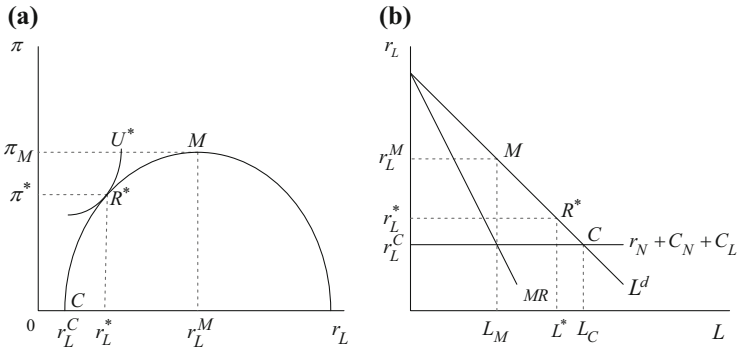

Fig. 9.1 A regulatory optimum

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 panel (a) of Fig. 9.1, the regulator desires a banking industry structure roughly midway between the extremes of perfect competition and monopoly. Without regulation, at a point near R^*, there is little incentive for either interest group to strive for establishment of a regulatory regime. At or near point M, however, consumers and elected officials 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 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

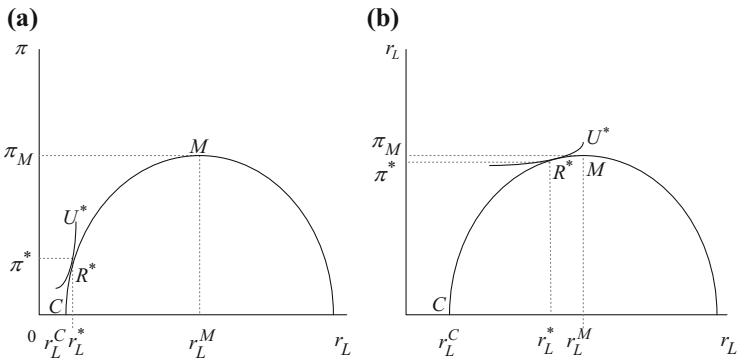

Fig. 9.2 Differences in regulator preferences

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).

Heinemann and Schüler (2004) find little support in data from more than a hundred countries in favor of the public interest hypothesis for bank regulation and conclude that private interests appear more likely to have prevailed. Rosenbluth and Schaap (2003) provide evidence from 22 nations suggesting that differences in electoral structures help to determine which type of bank regulatory regime, 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. Baxter (2011) argues that the threats posed by financial regulatory capture to the interests of taxpayers are sufficiently large and unaddressed by the Dodd-Frank Act and related reform efforts that additional and substantial regulatory restructurings should be contemplated.

As discussed in Chap. 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 bank regulators are particularly susceptible to 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 to 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.

Boyer and Ponce (2012) evaluate regulatory capture within an analysis of divisions of functions among different regulators. In their model, a "financial

stability committee" not dissimilar from the Financial Stability Oversight Committee established by the U.S. Dodd-Frank Act relies on either one or two bank supervisors who interpret signals about banks' prospects. The committee extends a supervisory "contract" to a regulator that includes an information-contingent menu of rules for banks and compensation schedule for supervisors—where the source of information is the signal extraction by the supervisor(s). Public-interested supervisors report fully information deciphered from observable signals, whereas self-interested supervisors may seek regulatory capture as a source of side payments from supervised banks in return for favorable reports to the committee.

Boyer and Ponce find that when the committee relies on a one supervisor, self-interest and the potential for regulatory capture create a distortionary spillover. The side payment offered by banks to the supervisor depends on the magnitude of the gain to banks and the supervisor from colluding. Hence, the committee must reduce the potential gain from collusion between the banks and supervisor by tightening regulatory rules. In contrast, when the financial stability committee utilizes two independently functioning, self-interested supervisors blocking the transmission of at least one correct signal to the committee would require collusive side payments from banks to both supervisors. In this case, the expected potential for supervisory capture decreases, and regulatory rules need not be as tight. Finally, if a supervisor does not observe the other supervisor's report—that is, in effect operates as a junior, subordinate supervising agent reporting to both the committee and the senior supervisor, the magnitudes of the distortions are unambiguously greater than in the case of independent supervisors. A regulatory structure with multiple independent supervisors, they conclude, is preferable to a structure in which one supervisor is subordinate to others. Nevertheless, multi-regulator structures yield better results than a single-regulator structure.

Hakenes and Schnabel (2014) propose regulatory capture as a possible explanation for poor banking outcomes and provide a model in which complex banks effectively "bowl over" insufficiently sophisticated regulators. In their model, regulators concerned about maintaining their public reputations approve continuing operations of such institutions rather than admit their inability to assess their risks of failure. In equilibrium, therefore, regulation is weak, and complex banks engage in risky activities.

The literature on regulatory capture—see Dal Bó (2006) for a review—suggests that one means of providing side payments to captured regulators is through a "revolving door," in which individuals alternate between employment with regulated firms and with supervisory agencies. Indeed, Bond and Glode's (2014) model of a labor market for individuals possessing who function either as bankers or bank regulators offers predictions consistent with a revolving door. Consistent with the basic story of Hakenes and Schabel (2014), in the equilibrium of the Bond-Glode model, the fact that the average remaining regulatory workers are less skilled than those at banks, bank misbehavior increases during financial booms.

Shive and Forster (2017) provide empirical evidence on the nature of the revolving door between supervisory agencies and banks. They find strong evidence that the door does indeed revolve. Regulated financial firms are twice as likely to

hire former regulators than are unregulated firms. Individuals hired away from supervisory agencies also receive a salary premium. On average, the regulated financial firm that hire former agency employees earn announcement returns, Shive and Forster seek to differentiate between the hypothesis that the regulated firms gain from consequent preferential treatment and insider information about supervisors' methods versus the hypothesis that agencies provide "schooling" for employees that those hired by regulated firms utilize to bring about risk reductions for the banks. Based in part on an empirical finding that downside risk for regulated financial firms typically follow the hiring of former employees of the firms' supervisory agencies, Shive and Forster choose the latter hypothesis as most supported by their analysis. As noted toward the end of this chapter, however, this empirical result could provide a measure of support for an alternative model of regulatory capture.

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. In addition, the analyses of Chaps. 3 and 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. Furthermore, the analysis of the effects of U.S. branching deregulation by Beck et al. (2010) suggests the potential for social-welfare improvements from less, rather than more, banking regulation, at least across some dimensions.

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. It is conceivable, therefore 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 inßuence its supervisory actions.

There is a small literature on how a regulator's preferences shape its actions when confronted with situations in which they must choose between engaging in forbearanceÑpermitting 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]. Some studies focus on the issues faced by a cost-minimizing deposit insurer/supervisor. 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 when 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 the monitoring reputation of a regulator charged with monitoring and closing banks can either respond to a sub-optimal Þrst-period risk choice by requesting a better risk choice the next period or closing the bank. They Þnd 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 regulatory 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 closure policies. They focus on a trade-off between the inßuence of policies on bank risk-taking and a social cost of closure resulting in the loss of intermediation services. They Þnd that 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].

Clark and Jokung (2014) offer a theoretical analysis of the credibility of regulatory interventions on the part of bank supervisors. Their analysis indicates that in the presence of regulatory costs and imperfect credibility, no universally optimal policy rule for interventions exists. Instead, the regulator's degree of risk aversion determines the choice.

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) 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.

In an analysis of 1993–2012 U.S. bank supervisory examination data, Rezende (2014) compares ratings with new regulatory supervisors after banks that have engaged in charter flips with the ratings they had previously received from their old supervisors. Rezende concludes that there a substantial positive effect of switching chargers on institutions' post-charter-flip examination rating—that is, banks can engage in beneficial regulatory arbitrage via charter flipping. He notes that this result indicates the possibility that regulators compete for clientele by awarding higher ratings to banks that switch charters.

In a study comparing the tightness of enforcement of banking regulations across U.S. federal and state regulators, Agarwal et al. (2014) find evidence of differences in preference weights of bank regulators. These differences, they argue, help to explain divergences that they document in the extent to which the same regulations are enforced by federal versus state authorities. Of course, regulators likely also

face differential resource constraints. Budget constraints and resource allocations of bank regulators have received relatively little attention from economic researchers. An important exception is Eisenbach et al. (2016), who provide analysis of supervisory resource distributions and document the extent to which resources of regulators are allocated to the largest and most complex banks.

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 induce a typical bank's expected net income to be higher if it takes the non-risky action while satisfying the regulator's budget constraint, inclusive of examination costs. If the regulator incurs an examination cost 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.

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: (1) 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 constraint. The other constraint is the incentive-compatibility constraint, IC. On or above and to the left 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.

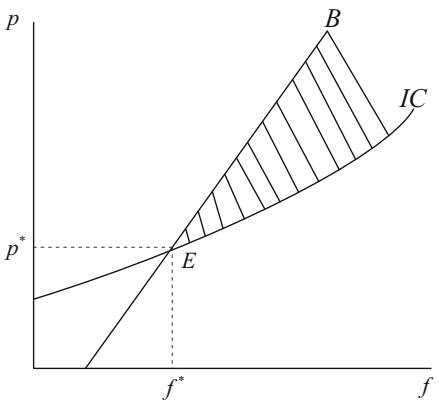

Fig. 9.3 A single regulator's supervisory policy choice

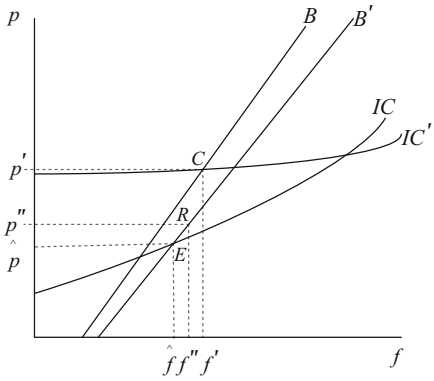

Fig. 9.4 Two types of banks

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, 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 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 the Basel standards consists of regulators possessing whatever degree of discretion is granted by national governments.

A Supervisory Race to the Bottom?

As discussed in Chap. 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.

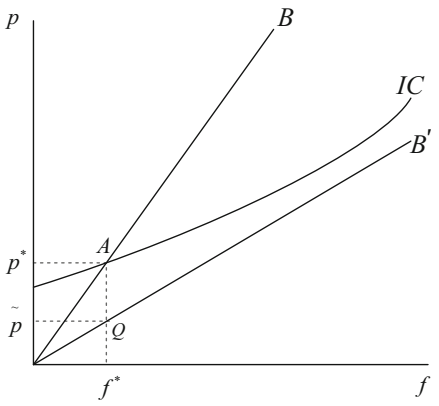

Fig. 9.5 The case of unconsolidated regulatory budget constraints

Figure 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. 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.

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 $p = f = 0$, represent a feasible outcome of a bidding war involving sufficiently non-conservative regulators. Thus, a true regulatory race to the bottom can take place. This analysis suggests that a race

Table 9.1 Bank regulator architecture: economic analysis versus the U.S. Dodd-Frank Act

Regulatory issue	Recommendation of economic analysis	Source of analysis	Dodd-Frank/FSOC stance on issue	Consistent with economic analysis?
Discretion available to regulators	Minimal supervisory discretion	Martimort (1999)	Wide latitude for discretion in "rulemaking," regulatory scope, and budgeting, particularly for FSOC and Consumer Financial Protection Bureau	No
Information used by regulators	Publicly available information	Tirole (1994), Laffont and Tirole (1991)	Both public information and inside information available only after the fact via Freedom of Information Act	No
Role of market-based data	Express and pursue goals in terms of market-based data	Helm (2006), Levine (2010)	Supervisory command and control using both market and inside information	No
Incentive-based contracts for regulatory supervisors	Incentive-based contractual arrangements appropriate to keep regulators focused on key goals	Samuel (2009)	Regulatory supervisors continue to operate within traditionally bureaucratic structure for performance evaluation	No
Horizontal centralization of regulation	*Pro*: Potential race to bottom, duplication cost inefficiencies *Con*: Specialization, competition gains, less supervisory flexibility	Esty and Geradin (2000), Reinicke (1994), Weinberg (2002) Rosen (2003), Dell'Ariccia and Marquez (2006), Masciandaro et al. (2013)	Reduced rivalry among commercial bank regulators left in place but operating within Financial Stability Oversight Council (FSOC) superstructure	Uncertain
Vertical integration of regulation	*Pro*: Deposit insurer duty reduces race-to-bottom potential *Con*: Increased corruption-capture potential *Mixed*: Bias toward regulatory forbearance depends on sizes of shocks	Weinberg (2002) Laffont and Martimort (1999), Esty and Geradin (2000), Easterbrook (2003) Kahn and Santos (2005)	More vertically integrated financial regulation via establishment of FSOC composed of regulators supervising different functions of financial firms	Uncertain

(continued)

Table 9.1 (continued)

Regulatory issue	Recommendation of economic analysis	Source of analysis	Dodd-Frank/FSOC stance on issue	Consistent with economic analysis?
Hierarchical predominance of certain regulatory supervisors	Potential for capture reduced if functional regulators have coequal status	Boyer and Ponce (2012)	Treasury secretary chairs FSOC; informational asymmetries among bank regulators	No

Source: Adapted from VanHoose (2013)

The Supervisory Review Process Pillar of the Basel Standards

Issues raised by the overlapping bank regulatory jurisdictions of the Federal Reserve, the Federal Deposit Insurance Corporation, and the Office of the Comptroller Currency are not unique to the United States. In Europe and other locales, regulatory jurisdictions are becoming more overlapped.

The latter part of Chap. 10 will address whether this fact helps to justify whether national governments should seek to bring their regulatory systems into greater harmony. In fact, however, they partially have done so within the supervisory review process pillar of the Basel regulatory framework. There are four key principles of the Basel 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)

In addition, "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 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 the results of this process (p. 209).

Left unstated is how and when supervisors should review 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,

> . . .[t]here 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 sufficiently brief 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 the Basel standards, 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 the Basel standards. For instance, Hamalainen et al. (2003) critique the Basel framework 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. Nevertheless, 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 institutions, resulting in regulatory forbearance in the face of weaknesses of troubled institutions.

Are regulatory actions triggered by specific events preferred over discretion? A considerable literature evaluates 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.

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 is 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 (2011) 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. John Walter (2004) provides a description of this process. 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—although Chernykh and Cole (2015) conclude that supervisors appear uncertain about appropriate tripwires from among these measures.

Do the prompt corrective action provisions of the 1991 FDIC Improvement Act yield a well defined bank closure process? From a theoretical perspective, Freixas and Parigi (2007) analyze 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. Harris and Raviv (2014) suggest that effective prompt-corrective-action policies can require a combined carrot-stick approach in which regulators reward banks for reporting problems while constraining their activities. Henderson and Tung (2012) propose rewarding bank regulators and supervisors for undertaking prompt corrective actions that prevent social losses.

Empirical evidence is mixed. Based on a cointegration analysis of quarterly 1984–2011 banking data, Gwilym et al. (2013) conclude that the 1991 adoption of prompt-corrective-action provisions reduced banks' default risks. In a comparison between the late 1980s-early 1990s savings bank meltdown and the crisis that began in the late 2000s, Balla et al. (2015) also find evidence that the prompt-corrective-action provisions helped to limit failures in the latter case. FDIC losses were proportionately higher, however, even though the effects of the two crises on the classes of banks examined by Balla et al. otherwise were similar, perhaps because, as discussed by Cowan and Salotti (2015), the FDIC had to subsidize acquisitions of failed banks to induce bids. Of course, as emphasized by Barth et al. (2014a, b), the utilization of prompt-corrective-action tripwires during the crisis episode was limited even though, as documented by Cole and White (2012) and Gaul and Palvia (2013), standard regulatory measures have provided clear prospective signals weak bank performances. Such banks have tended, when they failed, as shown by DeYoung and Torna (2013), to exhibit risky choices regarding within their basic lines of business as well as less-traditional activities. Finally, Loveland (2016) finds a failure of the guidelines to generate prompt action, with regulatory forbearance stretching to 18 months.

The recent crisis episode revealed that failure of prompt correct action has broader implications. 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 prompt corrective action rules. As discussed by Taylor and Williams (2009) and Taylor (2009), the Federal Reserve's establishment of bank lending programs 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.

Mailath and Mester's (1994) analysis suggests that 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 it 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 that proper application of prompt-corrective-action tripwires could have done much to reduce the scope of the subprime meltdown.

Is There Really a Basel Supervisory Review Process?

The Basel framework's supervisory-review process permits considerable variation in regulatory standards. The literature discussed above 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 bank regulators in obviates any true process for governments adopting the Basel 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, the Basel framework does not require national regulators to specify clear rules that will guide their supervision procedures. It also is silent regarding rules for enforcement actions against poorly managed and troubled institutions. Thus, if owners and managers of banks will operate under internationally coordinated standards, the standards effectively will be enforced in whatever way supervisors see fit to enforce them, often in situations involving regulatory conflicts, as discussed by Wall and Eisenbeis (2000). Rather than ensuring that the Basel framework establishes supervisory standards at least as tough as those of the most stringent national regulator, its framers have granted 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 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 the Basel framework.

Indeed, Barth et al. (2004, 2006) provide evidence that provision of considerable scope for regulatory discretion likely is 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 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 the Basel 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. Based on a panel analysis of 1999–2007 data from more than 4000 banks from 72 nations by Barth et al. (2013) also find evidence of efficiency improvements from market-based regulation as well as a positive effect of stronger supervisory powers in nations in which regulators operate independently. In contrast to Pasiouras et al., however, they

conclude that tighter restraints on bank activities are associated with lower 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 1400 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 standards.

Assessing Banks' Costs of Basel Compliance: Economies of Regulation?

What are costs to banks of complying with the Basel framework? 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.

Rule-of-Thumb Estimates

Most references to the costs of complying with the Basel standards refer to rule-of-thumb estimates. References in the media speculate, for instance, that 10% or more of banks' information-technology spending has been allocated to Basel compliance since the early 2000s. Another rule of thumb [see, for instance, Hitt et al. (1998)] is that as much as 15% 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 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 7402 commercial banks then operating of about $290.3 billion. Conservatively assuming that 10% of these were costs related to information technology and that 10% of these expenditures were related to Basel compliance, the implication is that the estimated total cost Basel 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 that banks' information-technology-related expenses really are as much as 15% of their non-interest expenditures. Then applying the 10% rule of thumb for Basel compliance costs as a share of information technology spending to the 2006 data boosts the estimated overall Basel compliance cost—again, conservatively assumed to equal the discounted present value of the entire stream of Basel-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% and the share of the latter expenses related to Basel compliance is only 5%, then the point estimate for Basel costs drops to only about $725 million (about $98,000 per bank).

Estimates of Basel Compliance Costs Based on Survey Data

Instead of relying on rules of thumb, in the mid-2000s the Office of the Comptroller of the Currency (2006a, b) surveyed banks about their own estimates of costs in implementing with Basel rules. The OCC concluded that for the 65 *nationally chartered* banks it judged to be most likely to utilize the most advanced requirements in place at that time, the combined compliance costs (discounted present value for 2006) amounted to about $473 million. Across these 65 commercial banks, this estimate implied an average total compliance cost of just under $7.3 million per institution. The OCC further estimated that if *all* nationally chartered banks were to adopt those regulatory requirements, the combined compliance costs (again as a discounted present value of all costs as of 2006) would have been nearly $1.1 billion, or almost $680,000 per institution.

Of course, these OCC estimates applied only to nationally chartered banks, or just over 20% of the U.S. banking system (albeit a portion containing many of the largest U.S. banks) at that time. Thus, the low-end total compliance-cost estimate of $725 million yielded by a rule-of-thumb-style approach was almost certainly too low. Indeed, these OCC cost estimates for nationally chartered banks made 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 smaller banks required to satisfy less burdensome 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 most advanced 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 reported that there were 3246 banks with assets less than $100 million, 3662 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, suppose for the sake of conservatism that the OCC's 19 banks were the only banks that initiated full implementation of the most advanced regulatory standards and that the remaining 69 very large banks continued to utilize less burdensome 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"—Scale Advantages in Basel 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 the pre-Basel-III framework. Suppose that we were to use the midpoint of each classification for smaller 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 was the set required to use or opting to utilize the most advanced Basel rules then in place) 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 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.

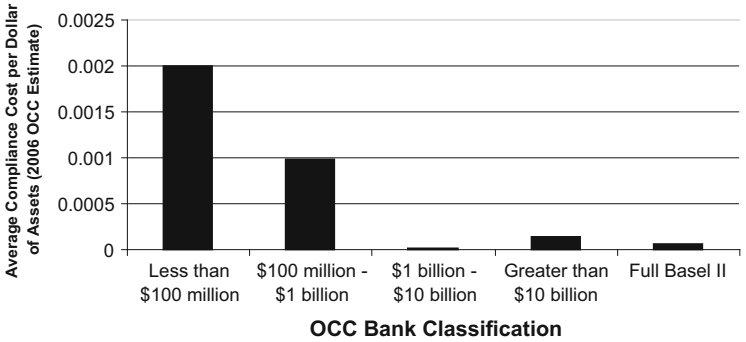

Fig. 9.6 Pre-Basel III average compliance costs by asset category

Second, among banks with more than $1 billion in assets, there may be compliance-cost diseconomies of scale for regional banks that increase their scale from the $1 billion to $10 billion size category to the greater than $10 billion size classification. Indeed, the OCC's estimates indicated that at some point beyond an asset size of $10 billion, a bank experienced lower average compliance costs by opting for advanced internal ratings-based regulation under the full Basel framework.

Figure 9.6 hints at the potential for significant average compliance-cost differentials. Is it possible to generalize about costs specific to Basel versus other non-interest expenses? On the one hand, the OCC (2006b) noted that the 19 nationally chartered banks using the most advanced Basel rules reported a per-bank average of total information technology expenses of $42 million that could be related to the Basel framework. 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% of banks' expenses on information technologies may be Basel-related. On the other hand, this "guesstimate" only applies to banks implementing the advanced Basel system. The OCC's compliance-cost estimates for smaller institutions suggested that smaller banks surely incurred compliance costs that were much less significant shares of their budgets for information technologies and other business lines.

To What Extent Do Basel Compliance Costs "Matter"?

What are the implications of Basel 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 implementation costs. The second part of the answer depends on an assessment of the on-going costs associated with complying with the new regulatory framework.

Total Implementation Costs

What does a discounted-present-value estimate of total Basel 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 compliance 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 imply the equivalent of a one-time 2 to 3.5% 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.

Effects of Basel Compliance on On-Going Expenses

Effects of Basel adherence on banks' on-going costs relate to aspects of compliance that impinge on their 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 likely on-going effects that compliance with the Basel framework on banks' 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, the OCC has provided an estimate of an annual expense of $2.4 million.

It is perhaps unsurprising that estimates of the on-going costs associated with Basel implementation are nearly nonexistent. After all, economists struggle to measure the relevant costs that banks already face under the existing regulatory framework. Evaluating the on-going burden of adherence to the Basel framework compliance clearly would require collection of detailed bank-level compliance-expense data.

On-going implementation of Basel framework 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 costs to society generated by capital-regulation-induced liquidity reductions is 0.1% of total consumption, or, based on 2006 data and in 2006 dollars, an annual decrease in consumption somewhat below $10 billion. Thus, while Basel 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 compliance for the banking system.

Bank Regulation and Endogenous Fixed Costs

As discussed in Chap. 3, 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, 2000) 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, Haldane (2012) suggests that the ramped-up regulatory complexity in the successive waves of Basel capital standards may generating substantial regulatory and compliance costs that increasingly 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% 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 successive waves of implementation of Basel standards. It is arguable that these 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 the Basel framework 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

regulator ensures attaining a maximum feasible level of "limit profits" (Geroski 1995). As in the Belleflame-Peitz rendition of Sutton's model, however, an incumbent bank must incur an exogenous entry cost, e. Furthermore, to attain the services of a captured regulator, each bank must provide the regulator with a fixed side payment, k.

Attaining a per-bank level of profits just sufficient to deter entry by a potential entrant bank requires the regulator to set u to satisfy the no-entry/positive-variable-profits constraint, $e + k \geq \pi^* = \frac{M\gamma y}{n_{reg}^2} - C(u_{reg}) > 0$, or $C(u_{reg}) \geq \frac{M\gamma y}{n_{reg}^2} - e - k$. That is, the regulator must set a minimum quality standard u_{reg} sufficiently high to discourage entry but maintain positive economic profits flowing to incumbents, which would entail making the inequality in the above expression an equality to just make the prospective entrant indifferent about entering the industry. Hence, in the hypothesized regulatory environment, the captured regulator selects the minimum quality standard in advance, taking into account the cost of entry and post-Cournot-rivalry variable profits in order to preserve the current number of bank incumbents, who then determine their outputs, the market loan rate, and profits—which again are just sufficiently positive to discourage any entry by non-incumbents. In the case of Belleflame and Peitz's endogenous-sunk-cost parameterization, $C(u) = \alpha u^{\beta}$, this implies that the appropriate minimum quality standard is $u_{reg} = \left(\frac{M\gamma y}{\alpha n_{reg}^2} - \frac{e+k}{\alpha}\right)^{1/\beta}$. Thus, the regulator subjects the banking industry to the requirement that the quality of the service flow from lending at least attain this value.

To understand the implications of this regulatory-sunk-cost model, take a look at Fig. 9.7, which is drawn in a format analogous to Figs. 9.1 and 9.2. In Fig 9.7, each variable $\bar{x}$ denotes the value in the absence of regulation and hence with exogenous sunk fixed costs for banks. Thus, without regulation, each bank's share of unregulated variable profits as a function of the loan rate is given by $\frac{\overline{\Pi}}{\bar{n}}$. The equilibrium flow of profits that just covers the exogenous entry cost e is at the Cournot market loan rate $\bar{r}_L$, at which the exogenous quality level is $\bar{u}$. Note that given a constant value of e, if an external shock were to cause variations in the typical incumbent's profit share, a change in the value of $\overline{\lambda}$ would bring about a variation in the value of $\bar{r}_L$. During the period of adjustment to new entrants, each bank's profit share also would change, potentially negatively over at least a portion of the adjustment period, until re-attainment of a new equilibrium following entry or exit of new banks.

Under regulation, the minimum quality standard established by the captured regulator would be the value of u_{reg} noted in the prior paragraph. At the corresponding value of the market loan rate, r_L^{reg}, each bank's share of variable profits is sufficient to prevent further entry. If the regulator maximized utility, given the regulator's underlying preferences regarding the bank's profit share and the loan rate, then the required side payment to compensate the regulator for delivering each bank's preferred outcomes is the value of k in the figure. In contrast to the unregulated case, an external shock generates an endogenous adjustment on the

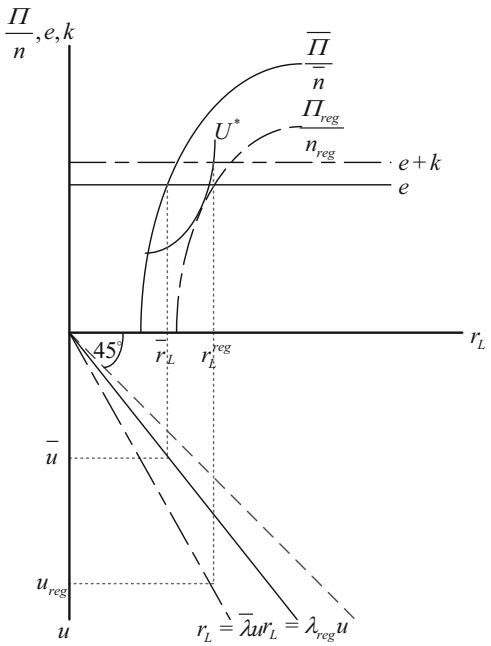

Fig. 9.7 Regulatory endogenous sunk fixed costs with a captured regulator

part of the regulator, which would respond by adjusting the value of u_{reg} as necessary to leave r_L^{reg}, and hence each bank's variable profits, unaffected. Consequently, establishment of minimum quality regulation administered by a compensated, captured regulator would yield stability of each bank's profits in the face of exogenous market shocks.

In the static equilibrium outcomes that apply to the unregulated and regulated environments in Fig. 9.7, net limit profits adjust to equality with the exogenous entry cost in the former case and with the sum of the entry cost and regulatory side payment in the latter case. During the adjustment to external shocks, profits vary in the former case but not in the latter case. Thus, the latter environment is preferable for risk-averse banks, whose owners benefit from the resulting decrease in downside profit risk.

Note that a reduction in downside risk is a key empirical finding obtained by Shive and Forster (2017) in their analysis of revolving-door flows of employees between supervisory agencies and regulated financial firms. An implicit promise by these firms of future employment to the regulator could, in principle, be an element of the side payment k in the above model, and the model implies that regulated banks' expense on such a side-payment would indeed yield a downside-risk reduction.

This example shows that 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. Given that the entry-signaling effects of endogenous regulatory sunk fixed costs typically are difficult to disentangle from exogenous elements influencing bank entry decisions—see Adams and Gramlich (2016)—evaluating empirically the predictions of the above theory would prove to be a challenging task.

Nevertheless, until the likely role of endogenous sunk fixed *regulatory* costs in determining the equilibrium structure of the banking industry has been resolved, policymakers should not accept at face value the notion that banking may be regarded as essentially a type of "natural oligopoly." Instead, as discussed by Saunders (2015), a tightly regulated environment may have transformed banks into institutions akin to public utilities subject to rate-of-return regulation. As a consequence, banks' profits likely will be more stable and limited to levels sufficient to provide a quiet life to bank owners while protecting them from the threat of entry by competitors.

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.

Acharya, Sankarshan, and Jean-Frncois Dreyfus. 1989. Optimal bank reorganization policies and the pricing of federal deposit insurance. *Journal of Finance* 44: 1313–1333.

Adams, Robert, and Jacob Gramlich. 2016. Where are all the new banks? The role of regulatory burden in new charter creation. *Review of Industrial Organization* 48: 181–208.

Agarwal, Sumit, David Lucca, Amit Seru, and Francesco Trebbi. 2014. Inconsistent regulators: Evidence from banking. *Quarterly Journal of Economics* 129: 889–938.

Agur, Itai. 2012. Multiple bank regulators and risk taking. *Journal of Financial Stability* 9: 259–268.

Allen, Linda, and Anthony Saunders. 1993. Forbearance and valuation of deposit insurance as a callable put. *Journal of Banking and Finance* 17: 629–643.

American Bar Association. 2007. Bank Mergers and Acquisitions Handbook. Section on Antitrust Law.

Balla, Eliana, Edward Simpson Prescott, and John Walter. 2015. Did the financial reforms of the early 1990s fail? A comparison of bank failures and FDIC losses in the 1986-92 and 2007-13 periods. Federal Reserve Bank of Richmond Working Paper 15-05, May 15.

Barth, James. 1991. *The Great Savings and Loan Debacle*. Washington, DC: American Enterprise Institute Press.

Barth, James, Gerard Caprio Jr, and Ross Levine. 2004. Bank regulation and supervision: What works best? *Journal of Financial Intermediation* 13: 205–248.

Barth, James, Gerard Caprio Jr, and Ross Levine. 2006. *Rethinking Bank Regulation*. Cambridge, UK: Cambridge University 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, Chen Lin, Yue Ma, Jesús Seade, and Frank Song. 2013. Do bank regulation, supervision, and monitoring enhance or impede bank efficiency? *Journal of Banking and Finance* 37: 2879–2892.

Barth, James, Gerard Caprio, and Ross Levine. 2014a. Misdiagnosis: Incomplete cures for financial regulatory failures. Milken Institute Center for Financial Markets, October.

Rochet, Jean-Charles. 2008. Why are there so many banking crises? In *Why Are There So Many Banking Crises: The Politics and Policy of Bank Regulation*, ed. Jean-Charles Rochet, 22–34. Princeton, NJ: Princeton University Press.

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.

Samuel, Andrew. 2009. Preemptive collusion among corruptible law enforcers. *Journal of Economic Behavior and Organization* 71: 441–450.

Santos, Joao. 2001. Bank capital regulation in contemporary banking theory: A review of the literature. *Financial Markets, Institutions, and Instruments* 10: 41–84.

Saunders, Anthony. 2015. Is Basel turning banks into public utilities? *Journal of Financial Perspectives* 3(Issue 1), March.

Shadow Financial Regulatory Committee. 2001. Statement on the Basel Committee's revised capital accord proposal, Statement No. 169, February 26.

Shim, Ilhyock. 2011. Dynamic prudential regulation: Is prompt corrective action optimal? *Journal of Money, Credit, and Banking* 43: 1625–1661.

Shive, Sophie, and Margaret Forster. 2017. The revolving door for financial regulators. *Review of Finance*.

Sleet, Christopher, and Bruce Smith. 2000. Deposit insurance and lender-of-last-resort functions. *Journal of Money, Credit, and Banking* 32: 518–575.

Stigler, George. 1971. The theory of economic regulation. *Bell Journal of Economics and Management Science* 2: 3–21.

Stiroh, Kevin, and Philip Strahan. 2003. Competitive dynamics of deregulation: Evidence from U.S. banking. *Journal of Money, Credit, and Banking* 35: 801–828.

Sutton, John. 1991. *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*. Cambridge, MA: MIT Press.

Sutton, John. 2000. *Technology and Market Structure: Theory and History, 2001*. Cambridge, MA: MIT Press.

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. Manuscript, University of Montreal.

Tirole, Jean. 1994. The internal organization of government. *Oxford Economic Papers* 46: 1–29.

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.

Van den Heuvel, Skander. 2008. The welfare cost of bank capital requirements. *Journal of Monetary Economics* 55: 298–320.

VanHoose, David. 2013. Evaluating the centralized-layers approach to U.S. federal financial regulation. Mercatus Center, George Mason University, July 15.

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 *Handbook of Financial Intermediation and Banking*, ed. Anjan Thakor and Arnoud Boot, 441–447. Amsterdam: Elsevier.

Wall, Larry, and Robert Eisenbeis. 2000. Financial regulatory structure and the resolution of conflicting goals. *Journal of Financial Services Research* 17: 223–245.

Walter, John. 2004. Closing troubled banks: How the process works. *Federal Reserve Bank of Richmond Economic Quarterly* 90: 51–67.

Weinberg, John. 2002. Competition among bank regulators. *Federal Reserve Bank of Richmond Quarterly* 88 (Fall): 19–36.
Whalen, Gary. 2002. Charter flips by national banks. Office of the Comptroller of the Currency Economic and Policy Analysis Working Paper 2002-1, June.
Whalen, Gary. 2008. Why do de novo banks choose a national charter? Working Paper, Office of the Comptroller of the Currency, December.
White, Lawrence. 1991. *The S&L Debacle: Public Policy Lessons for Bank and Thrift Regulation*. Oxford, UK: Oxford University Press.
Winston, Clifford. 2006. *Government Failure Versus Market Failure*. Washington, DC: AEI Brookings.

Chapter 10
Macroprudential Regulation and International Policy Coordination

In the aftermath of the 2007–2009 global financial meltdown and economic downturn and growth slowdown that followed, most nations' governments mounted substantial efforts to rein in risks of future system-wide crises. One focus of these efforts has been the contemplation of and utilization of so-called macroprudential policies to contain sources of systemic risk. The other area of emphasis has been consideration of and implementation of broadened forms of international regulatory policy coordination. This chapter considers both of these issues.

Systemic Risk

There is nothing at all novel about the concept of systemic risk, which has occupied the attention of central bank officials, bank regulators, and financial economists for more than four decades (see, for instance, VanHoose and Sellon 1989). What is new is governmental interest in explicit regulation of systemic risks—even though old issues relating to defining and measuring systemic risks remain not completely resolved.

Defining Systemic Risk

Systemic risk is a slippery concept, and VanHoose (2011a) provides a list of at least a dozen definitions. With the exceptions of Barth et al. (2014) and Kane (2010a, b)—who as discussed below place more emphasis on government's role as a potential source—the alternative definitions mainly tend to coalesce around two traditional notions of systemic risk discussed by Furfine (2003). One form of systemic risk identified by Furfine is "the risk that failure of one or a small number of institutions will be transmitted to others due to explicit financial linkages across

© Springer-Verlag GmbH Germany 2017

D. VanHoose, *The Industrial Organization of Banking*,
DOI 10.1007/978-3-662-54326-9_10

institutions." This view of systemic risk emphasizes supply-side channels of immediate interest to central banks concerned with the functioning of interbank clearings of daily transactions arising from interbank loans, securities exchanges, and foreign currency transfers. For instance, failure of a transaction to settle between two banks can result in a third bank failing to receive a promised payment from one of the first two banks, which in turn can generate settlement failures down a line of institutions. Thus, this notion of systemic risk views banks as making conscious decisions to establish systemic linkages that, in the event of settlement breakdowns among a subset of institutions, can spill over to adversely affect banks' customers, which, ex-post, are passive, essentially captive agents.

The second type of systemic risk that Furfine identifies is "the risk that some financial shock causes a set of markets or institutions to simultaneously fail to function efficiently." That is, financial markets experience a generalized collapse, presumably measured in terms of evaporating liquidity or solvency. As discussed in Chap. 5, one view on the mechanism for such a collapse is provided by theories in the mold of Diamond and Dybvig (1983), which thereby emphasizes demand-side sources of systemic risks emanating from proactive decisions by bank depositors. These depositor choices generate essentially passive responses—such as suspensions of convertibility of deposits or failure—on the part of banks that are essentially captives to consequent "runs." An alternative perspective on the sources of collective collapses focuses instead on banks' own decisions. This view emphasizes "herding behavior," which makes banks more susceptible to runs [see, for instance Acharya and Yorulmazer (2007, 2008)].

Irrespective of which of the two traditional views on systemic risk that one chooses as more appealing, there are three key *private-market*-related factors that potentially contribute to systemic risk. The first of these is *fractional reserve banking*. Commercial banks and many other types of depository firms hold reserves that are less than the total amount of customer deposits and other relatively short-maturity liabilities. Hence, each institution is always potentially susceptible to a positive net outflow of funds arising from its interactions with its depositors and other customers, thereby resulting in at least a temporary liquidity shortfall that generates susceptibility to runs. The second element is the potential for banks to engage in *correlated strategies*, or herding behavior. As discussed by Wagner (2008, 2010), a consequence is that traditional diversification at the level of individual institutions does not necessarily reduce systemic risk, because if multiple institutions choose similar diversification strategies, herding problems magnify.

The third contributor to systemic risk is *network externalities*, which in general arise when the magnitudes of the benefits or costs that people or firms receive or incur through their individual transactions in a market depend on how many others also engage in market transactions. In banking markets, network externalities and associated market feedback effects most obviously arise in the course of depository institutions' operations as payment intermediaries [see, for instance, Chakravorti (2000) and Eisenberg and Noe (2001)]. In fulfilling their payment-intermediation role, banking institutions participate in interlinked payment systems in which risk realizations between two parties can spill over to affect other parties implying

potential spillover, or systemic risks [see Wagner (2009) for a recent review]. Payment-system-related network externalities are particularly easy to visualize, but financial firms such as banks also experience network externalities in other contexts. For example, in interbank lending markets such as the federal funds market and derivatives markets, a default by one institution on a debt to another institution can potentially impinge on the latter institution's capability to honor debts owed to other institutions. In principle, the result could be a cascade of failures to honor obligations, resulting in realizations of systemic risks in the market [see, for example, Cocco et al. (2009), Fecht and Grüner (2007), Bliss and Kaufman (2006), Iori et al. (2006), and Freixas et al. (2000)].

Measuring Systemic Risk

In years past, economists have tended to rely on aggregate prudential ratios, such as aggregate capital ratios to evaluate systemic soundness of the banking industry. Such ratios are commonly used as well in the construction of possible early warning systems for banking and financial crises [see Barrell et al. (2010) for a recent example and Davis and Karim (2008) and Evans et al. (2000) for reviews]. Čihák and Schaeck (2010), however, document that even though such ratios sometimes have some informational content, they also suffer from some potentially significant difficulties. Among these difficulties with aggregate ratios are failures in identifying individual problem institutions and an over-reliance on backward-looking regulatory data instead of forward-looking market information.

group emphasize measuring exposures to contagious spillover effects of shocks experienced by "systemically important" institutions onto other institutions via linkages arising within interbank payment or funding networks.

Doing Something About Systemic Risk: Macroprudential Policy

By their nature, systemic risks arise from interactions among institutions, a fact that leads many observers to conclude that systemic risks cannot be addressed by traditional bank supervision and regulation. Wyplosz (2009), for example, criticizes traditional "microprudential regulation," which "concerns itself with the stability of individual institutions" and examines "responses of an individual bank to exogenous risks." His assessment is that traditional regulation "does not incorporate endogenous risk" and "ignores the systemic importance of individual factors such as size, degree of leverage, and interconnectedness with the rest of the system."

In light of these points, Wyplosz argues, macroprudential regulation is required. According to Hanson et al. (2011), "[i]n the simplest terms, one can characterize the macroprudential approach to financial regulation as an *effort to control the social costs associated with excessive balance-sheet shrinkage on the part of multiple financial institutions hit with a common shock* [emphasis in original]." Given potentially significant social costs, however these risks are defined or measured, policymakers have expressed a strong intent to mitigate them [for example, Bernanke (2008)].

Combating systemic risks is not a new objective of regulators. Indeed, in recognition of the potential for traditional supervisory and regulatory policies to miss the mark with regards to systemic risks, active discussion of macroprudential regulation were already under way for a number of years prior to the 2007–2009 banking crisis. A good example is Borio (2003), but three decades ago the Federal Reserve struggled with the issue of how to reduce systemic risks in interbank markets arising from daylight overdrafts by banks in U.S. large-value payment systems. In fact, Clement (2010) traces the term "macroprudential" back much further, to a 1979 meeting of the Cooke Committee (a forerunner to the Basel Committee on Banking Supervision) a report that year by a working committee chaired by Alexander Lamfalussy at the Bank for International Settlements.

Implementation of Macroprudential Policy

If one accepts Wyplosz's argument and Borio's (2009) analogous conclusion that macroprudential regulation is "an old idea whose time has come" [also see Davis

and Karim (2009)], how might a macroprudential regulatory regime be structured and implemented?

Structuring Macroprudential Regulation

The first step in implementing macroprudential policy would be to identify the set of institutions upon which a macroprudential focus should be placed. Brunnermeier et al. (2009) offer a typical proposal for how to accomplish this step. They suggest (pp. 25–26) that systemic-risk measures can be applied *objectively* to classify institutions—not just banks but any "financial players" that are "subject to systemic risk"—into one of the following categories: (1) *individually systemic institutions* that are so large and interconnected that a national government would under circumstances regard them as too big to fail; (2) *systemic-as-part-of-a-herd institutions* that are individually small but tend to engage in highly correlated strategies; (3) *non-systemic large and not highly levered institutions* that tend to pursue generally uncorrelated strategies; and (4) *tiny institutions* that are too small to pose systemic risks and which also tend not to herd with other institutions. Brunnermeier et al. (2009) suggest that macroprudential supervision and regulatory policies should be directed toward institutions that objective systemic-risk measures classify within the first two categories. They argue that institutions falling within the latter two groups should be subjected only to traditional micro-based supervision and regulation.

As shown in Fig. 10.1, Brunnermeier et al. (2009) effectively propose leaving in place for all regulated financial institutions micro-based policies directed toward individual institutions in isolation. Macroprudential regulation would supplement the traditional set of supervision-and-regulation policies among the subset of institutions that systemic-risk measures—such as the U.S. Dodd-Frank Wall Street Reform and Protection Act of 2010—identify to be either within the individually systemic—greater than $50 billion in assets in the case of the Dodd-Frank Act—or systemic-as-part-of-a-herd categories.

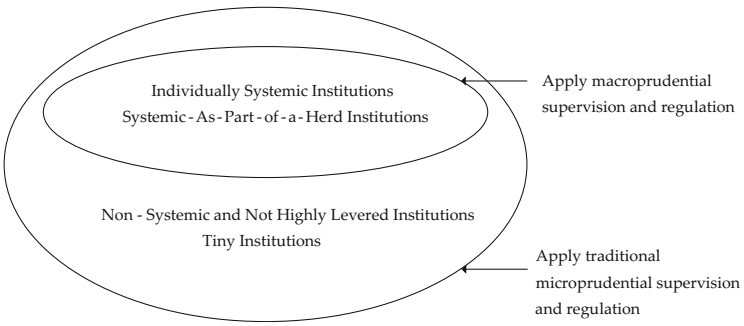

Fig. 10.1 The Brunnermeier-Crockett-Goodhart regulatory scheme

Because common shocks might arise in conjunction with cyclical economic and financial swings that influence the course of a bank's performance [see, for example, Albertazzi and Gambacorta (2009)], Persaud (2009) emphasizes the importance of a focus on establishing a system of macroprudential regulation that includes systems of automatic stabilizers. These might, he suggests, include upward and downward adjustments of capital adequacy regulations and liquidity buffer requirements during the course of business cycles. Persaud also recommends the use of mark-to-funding accounting and regulations concerning bank management compensation aimed at focusing bank managers' decision-making away from immediate-run horizons [see VanHoose (2011b, c) for a review of issues relating to regulation of bank management compensation].

Rochet (2010) argues that the current macroprudential emphasis on "systemically important institutions" is misplaced. In his view, macroprudential regulation should be focused on protecting *platforms, which i*n the language of the recent literature on two-sided markets are intermediaries that bring together different groups of "end users," such as interbank-credit clearing systems and payment clearing mechanisms that service two sides of transaction-based two-sided financial markets [see, for instance, Rochet and Tirole (2006, 2008)]. Specifically, Rochet (2010, pp. 272–274) suggests that macroprudential regulators should "guarantee the integrity of a precise list of financial markets and infrastructures that are deemed 'vital': interbank. . .markets, money markets, and some derivative markets and large-value payment systems." Rochet proposes that these platforms "would only be directly accessible to a group of 'officially recognized financial institutions' that would have to comply with special regulatory requirements," which he proposes could be based on "market best practice[s] put in place by private central clearing platforms and would be directly supervised by the central bank" and that "the central bank would choose who belongs to the club and who does not." Thus, Rochet would refocus macroprudential regulation onto clubs of government-selected institutions interacting on platforms with focuses circumscribed by supervisory authorities, an approach that he argues would, among other things, "eliminate the rationale for splitting or downsizing banks."

Additional Macroprudential Policy Instruments

What policy tools should regulators bring to bear to implement macroprudential supervision? The basic menu of instruments begins with existing tools. For instance, consider macroprudential issues discussed by Kashyap et al. (2010), who emphasize the problem of "fire sales" [see von Peter (2009) and Caballero and Simsek (2013)] involving forced sales of relatively illiquid assets at prices dislocated from originally perceived fundamentals. Such fire sales, they argue, arise because banks face incentives to gamble that they can avoid failure and postpone selling relatively illiquid assets. Consequently, banks will tend to try and sell these assets only when they are on the cusp of failure. Analysis of a general-equilibrium model leads Kashyap et al. to conclude that the existence of multiple spillover effects from fire sales requires coordinated settings of multiple macroprudential

policy tools—selected from a set that includes longstanding instruments such as capital requirements and margin requirements.

The most publicly visible element of macroprudential policy involves "stress tests," or integrated capital-adequacy modeling calibrated to alternative economic scenarios entailing potentially interconnected effects on banks' balance sheets [see, for example, Kretzschmar et al. (2010)]. As emphasized by Daníelsson (2008), however, such calibrated simulations rely on risk models that commonly yield results subject to significant inconsistencies.

A macroprudential tool that Kashyap et al. (2010) add to the traditional listing is liquidity requirements. Indeed, a macroprudential policy instrument recently added to the Basel regulatory toolkit is a liquidity-coverage-ratio (LCR) requirement, which requires banks to hold a sufficient ratio of a stock of liquid assets—vault cash, central bank reserves, most government securities, and some corporate and mortgage-backed securities—to a typical 30 days' worth of the bank's net cash outflow [see Liebmann and Peek (2015) for a detailed discussion]. Typical discussions of the LCR requirement follow Carlson et al. (2015) by citing the liquidity shortfalls that arose at a number of financial institutions during the 2007–2009 as demonstrating a "need" for such a regulation. In addition, Distinguin et al. (2013) utilize data from 781 U.S. and European banks to study the relationship between banks' capital and liquidity positions and conclude that the two are interconnected. In particular, they find evidence that, with the possible exception of small banks in certain circumstances, banks tend to reduce their capital ratios in response to greater illiquidity. Nevertheless, as Cetina and Gleason (2015) discuss, U.S. implementation of the LCR requirement created a quite complicated and difficult-to-interpret measure of high-quality liquid assets and of the official LCR. Furthermore, in practice the computation of a bank's LCR involves relating the stock current liquid assets to previous cash outflows. Balasubramanyan and VanHoose (2013) focus on this aspect of the LCR requirement and conclude that it introduces a new source of dynamic balance-sheet linkages that yields ambiguous implications: Although intertemporal paths of bank assets and liabilities tend to exhibit greater persistence, they also become more responsive to shocks—a point emphasized as well by Jacob and Munro (2016).

Special taxes on banks are another novel macroprudential instrument. Following the financial crisis, such taxes were proposed to address a variety of sources of systemic risks by inducing banks to cut back on potential over-expansions of particular assets or liabilities. As discussed by Devereux (2014), a number of European nations, including Germany and the United Kingdom have imposed various forms of such bank taxes in recent years. Although such taxes certainly give banks incentives to constrain expansions of directly taxed assets or liabilities, feedback effects onto banks' balance-sheet adjustments and other decisions need not fully support the safety-and-soundness objectives of the taxes. For instance, Dia and VanHoose (2017) show that an unintended effect of taxing bank lending to restrain potential bank credit expansion can is an alteration of the overall composition of bank lending away from monitored loans and a consequent reduction in aggregate credit quality.

Pitfalls of Macroprudential Regulation

There are several obvious difficulties associated with implementation of any macroprudential-supervision framework. In light of the preceding discussion, one glaring difficulty is remaining uncertainties about defining and measuring systemic risks that such policies are intended to mitigate.

White (2008) worries about problems with "acceptance" by national regulators, regulatory inertia in developing a will to act at the appropriate time, and conflicts with monetary policymaking. In addition, Born et al. (2010) raise concerns about the use of speeches and interviews by macroprudential regulators, which they find, in analysis of more than 1000 speeches, interviews, and formal Financial Stability Reports (FSRs) for 36 nations, tend to add to financial instabilities, whereas FSR releases generally are associated with reduced instabilities. Daníelsson et al. (2016) express doubts about regulators' capabilities to identify systemically risky institutions given that estimation errors alone can impede such efforts.

It is arguable, however, that there are three deeper pitfalls that have yet to be fully investigated in the literature on systemic risks and macroprudential regulation. The first of these is the threat of an enhanced likelihood of regulatory capture, a topic discussed in Chap. 9. Applying this economic approach—reviewed most recently by Veljanovski (2010)—to macroprudential regulation indicates that government officials and even regulated financial institutions might promote establishment of another financial regulatory superstructure to enable the pursuit of self-interested goals. Alternatively, even if a public-interest motivation for such a superstructure prevails, incentives to pursue self-interests ultimately could skew day-to-day supervision in favor of providing protections for regulated financial institutions. Indeed, in many instances it would be difficult to disentangle the motivations for certain policy actions. As discussed by VanHoose (2013), increased centralization of and interconnection among the layers of regulatory oversight across multiple supervisors provides a number of potential routes to a regulatory-capture outcome.

Another potential pitfall associated with macroprudential regulation, discussed by Barth et al. (2014) and Kane (2010a, b), is the danger of overreliance on perceived market failures to rationalize such regulation without due attention to governmental regulatory failings. Finally, most research on addressing prospective systemic-risk externalities jump to proposals for public solutions, such as macroprudential regulation, without contemplating either potential private contractual solutions or ways in which expanding the financial regulatory superstructure could contribute to externalities associated with systemic risks. We shall return to this issue later in the chapter.

International Regulatory Policy Coordination

Alongside the growing interest in macroprudential regulation has been an upswing in proposed and implemented mechanisms through which national financial regulators and supervisors might more fully coordinate their efforts. Since the earliest discussions of international regulatory coordination in the 1980s, the bias was toward the view that international financial policy coordination is always desirable (see, for instance, Folkerts-Landau, 1990). Only within the past few years have very many banking and finance researchers raised questions about this long-held assumption on the part of many national supervisory authorities. The remainder of this chapter considers both sides of this discussion.

Essentials of the Regulatory Policy Coordination Problem

To contemplate the basic implications for financial regulation that are likely to arise from applying the fundamental lessons of the international-policy-coordination literature, consider a very simple setting similar to that of Weinberg (2002) that was discussed in Chap. 9. As in his model, suppose that financial regulatory authorities in two jurisdictions confront clienteles of regulated financial firms that can opt to switch from one jurisdiction to another via business reorganizations or perhaps just by obtaining charters from the other authority. Recent evidence on such international forms of such regulatory generally supports this assumption. Temesvary's (2015) study of 2003–2013 data on U.S. banks' activities in 82 nations concludes that U.S. banks are substantially more likely to choose to operate in countries with weaker regulatory requirements and avoid activities in nations with tighter rules. Berrospide et al.'s (2016) analysis of 2000–2013 data on domestic and foreign banks operating in the United States indicates that foreign regulatory tightening boosts growth of U.S. lending by subsidiaries of foreign banks and that toughened foreign banking rules induces shifts in lending away from nations with toughened rules to the United States and other nations with unchanged regulations.

In the model that follows, each financial regulatory authority has only one policy

every possible setting, Fig. 10.2 has been drawn under the simplifying assumption that noncoordinated policymaking yields a regulatory race to the bottom: At point *N*, both authorities opt to set the probability of auditing their clienteles of financial firms at zero, so no audits occur and failure costs are at their maximum. Offsetting these costs for each authority are the facts that it incurs no auditing costs and that it has eliminated any incentive for members of its clientele to switch jurisdictions.

In contrast, point *E* arises under international financial regulatory coordination, under which the two authorities *jointly* choose π and π^* to minimize their *combined* loss. With quadratic policy loss functions, the solution to this policy problem turns out to require the indifference ellipses of the policymakers to be at a point of tangency with equal slopes. Given the assumptions about the nature of policy losses discussed above, coordinated regulatory policies yield higher auditing probabilities at point *E* than is true of noncoordinated policies at point *N*. The latter policies that yield point *N* would fail to take into account that if both authorities were to coordinate joint increases in their audit probabilities to the levels at point *E*, they would simultaneously reduce their failure losses while continuing to prevent erosion of their clienteles of regulated financial firms. Figure 10.2's structure presumes that these loss-reduction gains would be sufficient to more than offset the higher auditing costs incurred, hence reducing both authorities' policy losses. Hence, in the example depicted in Fig. 10.2, coordination "pays" for the two authorities and constitutes an "efficient" equilibrium from their points of view.

Point *C* illustrates the difficulty associated with attaining point *E*. Once the two regulatory authorities agree to set their audit probabilities at levels consistent with point *E*, the domestic authority can gain by reducing its own audit probability in order to reach point *C* on an indifference ellipse closer to its own bliss point. That is, the domestic authority can gain by cheating on the policy agreement—as noted by White (1994), much like a firm in a cartel can gain from cheating on a collusive agreement. (Alternatively, but not shown in Fig. 10.2, the foreign authority also stands to gain from cheating and reducing its own audit probability if the domestic authority honors its side of the agreement.) The mutually shared incentive to cheat means that in the absence of a means of ensuring commitment to policy coordination, point *E* will be difficult to maintain. If either authority yields to the temptation to reduce its national policy loss by cheating, the policymakers eventually will gravitate back toward point *N*'s regulatory-race-to-the-bottom outcome. (Point *S* in Fig. 10.2 is discussed below.)

Recent Research on International Coordination of Financial Regulatory Policies

A number of studies have sought to evaluate international coordination of financial regulatory policies. One of the most widely cited is that of Dell' Ariccia and Marquez (2006), although earlier independent work by Dalen and Olsen (2003)

touches on some of the issues that they consider, as does more recent work by Boyer and Kempf (2016) that touches as well on issues discussed in Chap. 9. In Dell' Ariccia and Marquez's model, each regulator in two countries seeks to maximize a weighted average of the profits of its nation's single bank and of that bank's monitoring effort. Given the imperfectly competitive structure in Dell' Ariccia and Marquez's model, if one of the banks cuts back on lending in either country's market, the equilibrium loan rate rises, which pushes up average revenues at both banks. A tighter risk-based capital requirement at a domestic country's bank induces a lending cutback that simultaneously boosts market loan rates and the bank's average revenues from lending and increases its return from monitoring. The unambiguous consequences are increased monitoring and a more stable bank portfolio—outcomes that Chap. 6 shows do not necessarily hold true in all possible banking environments. Nevertheless, tighter capital regulation raises the costs for the domestic bank sufficiently to reduce its profits on net. Not surprisingly, therefore, a regulator that places a greater weight on bank profits relative to portfolio stability in its objective function prefers a lower capital requirement.

When the regulators do not coordinate their policymaking in this model, the domestic regulator balances the effects on profits and portfolio stability at its nation's bank. It ignores the fact that a higher loan rate caused by toughened capital requirements on the domestic also generates an increase in average revenues and in the return from monitoring at the foreign bank, hence increasing its portfolio stability as well. Consequently, uncoordinated policymaking by the domestic regulator neglects the positive spillover that a domestic capital tightening has on foreign portfolio stability, and vice versa for noncoordinated policymaking on the part of the foreign regulator [a conclusion reached by Sin (2001) in an earlier study.] Dell' Ariccia and Marquez proceed to show that regulators with the same degrees of relative weighting on bank profits and portfolio stability preferences will choose lower capital requirements than would be chosen if the two regulators instead worked together to coordinate their choices. Within the context of their model, this conclusion is analogous to the regulatory race-to-the-bottom result under noncoordinated policies depicted in Fig. 10.2, with capital requirements replacing audit probabilities as instruments of regulatory policy. A related conclusion is that inducing two regulators with identical preferences to agree to coordinated regulation requires the commonly applied capital standard to be set higher than the regulators would have established independently. Tougher capital requirements are necessary because they assure mutual internalization of the positive spillover effects that capital regulation has on portfolio stability. In addition, as differentials in regulators' preference widen, the likelihood of reaching an agreement on policy coordination diminishes, to a point beyond whichno agreement is feasible. Finally, even if regulators have identical preferences, a subset of regulators will prefer to remain outside a policy-coordinating group. Formation of such a group yields benefits to regulators in countries both within and without, which those on the external margin regarding joining the group must balance against lost bank profits. For some regulators beyond that margin, the reduction in profits can be too

substantial in relation to the already freely obtained stability benefits from the coordinating group to justify joining it.

Eldridge et al. (2015) generalize the Dell' Ariccia-Marquez model by allowing for country size to vary. They argue that regulators in small countries know that the reactions of their client institutions to supervisory rules will have meager effects on global interest rates and hence are more likely to opt for lax supervision to benefit their domestic clients. Hence, their conclusion is that regulators in small nations have greater incentives to deviate from financial-supervisory coordination agreements.

In contrast to Dell' Arricia and Marquez, Morrison and White (2009) examine a setting in which there is no international competition among banking systems. Regulators in separate locales can provide operating licenses to either domestic- or foreign-based banks, however. Consequently, in Morrison and White's model, bank capital (or, as they note in a footnote, alternatively the labor of managers with differing capabilities) is mobile across borders while cross-border trade in banking services does not take place. Morrison and White examine a trade-off that arises when banking systems in different nations establish a regulatory level playing field across capital-standard jurisdictions. On the one hand, a level playing field prevents an international policy spillover in the form of a "cherry-picking externality" when jurisdictions are differentially supervised. In the latter setting, banks gain reputational benefits and looser capital requirements that lead to lower funding costs if they opt to apply for charters to operate in jurisdictions with the best reputations for high-quality regulatory monitoring. This fact enables regulators possessing the best reputations to grant charters to the strongest banks, leaving weaker banks to be supervised by regulators in other jurisdictions. On the other hand, nations can eliminate this cherry-picking spillover via establishment of a level playing field in which the common regulation imposed on banks in all jurisdictions corresponds to the requirement that the weakest regulator would implement. Such a common capital-regulation framework, however, creates regulatory inefficiencies in countries with the highest-quality regulators. Morrison and White find that on net, permitting multinational banking and long-run adoption—presumably via international policy coordination—of level playing fields are the least costly approaches for regulators to pursue.

Clearly, the desirability of international policy coordination depends crucially on the extent to which nations' regulators face symmetric policy incentives. Beck et al. (2013) evaluate this issue. The authors initially set up a simple closed-economy framework in which a regulator supervises banks with purely domestic operations with an exogenous probability of failure. They suppose that the regulator will intervene and liquidate a bank offering a one-unit payoff in the event of failure when the known payoff exceeds the expected return from permitting continued operations. Next, Beck et al. contemplate a setting with cross-border banking, in which a certain payoff if the bank is closed is split between domestic residents and residents of other nations if the bank is closed. They find that that intervention is more likely when the foreign-held share of equity in the bank is larger and the foreign-held shares of deposits issued by the bank and assets held by the bank are

smaller. On the one hand, a larger share of foreign stakeholders in domestic banks leads to a greater misalignment of interests of the regulator with the interests of shareholders, which makes intervention by the domestic regulator more likely. On the other hand, a smaller share of deposits held by foreign residents or a smaller share of credit granted to foreign residents imply that the burden of failure will be felt more by domestic residents and makes immediate intervention more likely.

The emphasis that Beck et al. place on their results is the potential for inefficient supervision of European banks engaged in cross-border operations. Their analysis has another implication, however. In a foreign country in which the gross return on assets, external cost of failure, and other parameters all differ at least slightly from those for the domestic nation, either Beck et al.'s implied value of the efficient foreign success probability threshold for foreign regulatory intervention nor that of the inefficient, foreign-centric threshold would typically correspond to those for the domestic nation. Whether or not the domestic or foreign regulators were interested in pursuing efficient policies these authorities would exhibit different preferences, ceteris paribus, regarding points at which to intervene and liquidate (or bail out) banks. Hence, reaching agreement on coordinating supervisory rules and actions could well prove to be a problematic undertaking. Kopecky and VanHoose (2012) reach an analogous conclusion regarding any supranational establishment of capital requirements aimed at inducing socially optimal monitoring of loan quality by banks.

Beck et al. suggest that a possible response to the problem of reaching a multilateral agreement might be for nations to entrust financial regulation to a supranational institution. After all, as D'Hulster (2011) discusses, if parties are convinced that full coordination of national policies is desirable, the natural conclusion is to impose some sort of centralized rules aimed at eliminating adverse information-sharing problems owing to bureaucratic conflicts of interest. Problems relating to information exchange among supervisory authorities is the key focus of Holthausen and Rønde (2004), 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.

Acharya (2003) provides one of the more interesting studies aimed at evaluating practical issues likely to impede financial-supervisory-coordination efforts. Acharya sets up a model in which identical banks within a nation make either "safe" or "risky" investments financed by deposits that all pay the same promised rate of interest. Equity capital acts as a buffer by reducing the threshold point of default, which results in a bailout ("forbearance") with a known probability. Banks confront countervailing incentives: (1) On the one hand, they have a risk-avoidance

incentive because greater risky investment raises the likelihood of losing their charter values; (2) on the other hand, banks have an incentive to undertake risky investments to maximize the value of their option to default and be bailed out with the known probability. The bank's optimal capital choice in Acharya's model trades off the reduction in likelihood of failure resulting from increased capital against the associated dilution cost of capital. Greater forbearance already buffers the bank against the loss of charter value, so banks react to increased forbearance by preferring less capital. From banks' point of view, therefore, greater forbearance and the privately optimal capital level are strategic substitutes.

A bank regulator determines a minimum required capital threshold to maximize the expected continuation value of all bank claims given the bank's privately undertaken investment allocation decision, subject to incentive-compatibility and participation (charter value must be sufficiently high) constraints. Thus, if the regulator raises the bailout probability, the bank engages in more risky investments, so the likelihood of bank default and hence the probability of bailouts occurring also increases. Consequently, from the regulator's perspective, the optimal minimal capital requirement and regulatory forbearance are strategic complements.

Acharya assumes that the banking markets of the two countries are integrated, so banks, borrowers, and depositors can engage in unhindered cross-border arbitrage. The greater degree of forbearance on the part of the foreign regulator enables foreign banks to opt for a larger share of risky investments. The result is a spillover effect operating through cross-border competition that reduces profit margins for domestic banks. To prevent an exit of domestic institutions from the industry—that is, to ensure satisfaction of the participation constraint that it faces—the domestic regulator ultimately will be pressured to raise its own degree of forbearance. Hence, a race to the bottom in regulatory standards eventually must occur, unless the domestic regulator is willing to abandon the participation constraint and observe an evaporation of the domestic banking system. As long as the foreign regulator has a greater willingness to forbear, analogous pressures will be placed on the domestic banking system and regulator whenever the same minimum capital standard binds both nations' banks.

Acharya offers two possible policy responses to prevent adverse international spillover effects across bank regulatory regimes. The first of these is full international coordination of all aspects of bank regulation—that is, engaging in cooperative determination of all regulatory instruments, taking into account jointly any existing cross-country heterogeneities, and aiming for mutual attainment of regulatory objectives. The other policy option is host-country-centered regulation, under which any banks seeking to operate within a nation's borders would have to satisfy that regulator's requirements. In effect, the host country's regulator would grant charters to operate within its borders only to those institutions that agreed to subject themselves to the host regulator's supervision.

Engineer et al. (2013) apply an Acharya-style analysis to examine regulatory competition involving EU nations' deposit-insurance limits during the 2007–2008 crisis. They argue that the ratcheting up of these limits by nations' governments—essentially a regulatory race to the bottom in settings of per-dollar insurance

premiums—was the outcome of a noncooperative game among authorities seeking to prevent destabilizing outflows of bank deposits.

Can Impediments to International Bank Regulatory Coordination Be Surmounted?

Acharya's contrasting policy recommendations suggest that regulators with differing preferences or facing asymmetric national financial environments could face a fundamental tension. If policy coordination is feasible, then such coordination promotes financial stability when financial institutions engage in cross-border competition under a coordinated system of transnational supervisory rules. But if coordination is not feasible, his results imply that nations can achieve financial stability objectives only if their regulators abandon efforts to maintain open international competition on a playing field characterized by coordinated regulatory policies.

The idea that international stability, openness, and policy objectives can prove to be mutually incompatible in practice is not new. International monetary economists have identified a "trilemma" arising from an inability to achieve simultaneous goals of a fixed exchange rate, mobile financial capital, and independent national monetary policies. Some financial economists have argued that an analogous financial trilemma also exists: an alleged impossibility of simultaneously attaining objectives of national financial stability, international financial integration, and independently formulated national financial regulatory policies. That is, under a financial trilemma, a nation can attain any pair of these goals: financial stability and international integration, financial stability and independently pursued financial policy making, or international integration and independent financial policies. A nation cannot, however, achieve all three objectives.

Schoenmaker (2011) provides a model motivating the concept of a financial trilemma, in which authorities in different jurisdictions contribute funds to address the failure of a bank via a "refunding." Such action yields both a benefit, B, and involves a net cost, C. The benefit of such action is distributed fractionally across countries, whose national authorities also contribute fractional amounts of the required failure refundings. Greater international integration spreads the benefits more thinly across countries, which reduces the incentive for national authorities to contribute and thereby reduces financial stability. Moving to a system of coordinated refundings allows for financial stability but removes the potential for independent financial policy choices. Reduced integration tends to centralize benefits, which increases national authorities' incentives to provide stabilizing refundings but naturally entails less financial integration. Hence a financial trilemma follows naturally in Schoenmaker's framework.

The logic of the trilemma appears to imply that commitment to financial stability yields a stark either-or choice between a requirement to coordinate international

regulatory policies and an objective of achieving international financial integration. Gaspar and Schinasi (2010) offer an argument to the contrary, the essential elements of which can be assessed within the context of Fig. 10.2 above. They note that an efficient combination of coordinated financial regulatory policies, point E in the figure, would be readily attainable in the absence of Coasean (1960) transaction costs. In the event that such costs exist, however—for instance, because of heterogeneities in regulators' preferences or differences in the sizes or structures of their economies and financial markets—the feasibility for bargaining among national supervisors to yield the efficient outcome E becomes more limited. They argue that this fact does not necessarily rule out the feasibility of policy outcomes that are more desirable than point N, the outcome that emerges when there is no policy coordination whatsoever. Generally speaking, *any* policy bargain in the setting considered in Fig. 10.2 that moves both financial policymakers' choices to the upper right from point N—that is, in the direction of a range of points roughly between B and B^*—will result in gains for both regulators.

One bargain that conceivably might yield such an outcome would be for the foreign regulator to offer to allow the domestic regulator to act as the "first mover" or "Stackelberg leader." Under this bargain, the domestic financial supervisor would be permitted to establish its own preferred examination probability, taking into account how the foreign regulator will respond. In Fig. 10.2, this arrangement would yield point S, at which the domestic regulator's indifference curve is tangent to the foreign supervisor's reaction or best-response function, which the domestic supervisor regards as a constraint. Thus, the domestic regulator is permitted to establish an examination probability π_S, which is lower than the efficient, fully coordinated policy setting π_E. Yet the policy setting π_S provides the domestic regulator with "political cover" that may enable it to implement the bargain successfully. The foreign supervisor, which would act as a "Stackelberg follower," then optimally responds with a policy setting π_S^* below the efficient level π_E^*. Allowing the domestic regulator to act as a leader yields an outcome at point S that is less preferable to the foreign supervisor than the most efficient, fully coordinated outcome at E. Nevertheless, point S is more preferable to the foreign supervisor than the race-to-the-bottom outcome at point N.

Gaspar and Schinasi survey key economic arguments advancing the idea that bargains with outcomes that are strictly superior, such as point S in Fig. 10.2, to a purely noncoordinated equilibrium, such as point N, can be maintained by parties, such as domestic and foreign financial supervisors. They point out that these arguments indicate that many such bargains that are preferable to the noncoordinated outcome can be feasibly implemented, even if they are inferior to the fully efficient, coordinated outcome. This conclusion arises when the static, one-shot interaction among two parties such as those in Fig. 10.2 is replaced with repeated interactions. When the two parties interact repeatedly, a fundamental condition for feasibility of bargains such as the possibility described above is that the parties must be sufficiently "patient." That is, they must place sufficiently heavy weight on future outcomes relative to current outcomes.

The first-mover-assignment bargain is an interesting potentially feasible outcome because when interactions are repeated, the Stackelberg-style leader-follower structure theoretically can perpetuate itself. Suppose that in Fig. 10.2, once the two supervisors establish an agreement for the domestic regulator to act as the leader, if the domestic authority cheats by setting a lower examination probability than π_S. As part of the bargain with the domestic regulator, the foreign supervisor can make clear that in response to such an event, in the following and all subsequent periods, it will revert to the noncooperative policy setting $\pi_N^* = 0$. As long as the domestic regulator places sufficient weight on attaining the point S outcomes in these future periods—which from its perspective are more desirable than point N outcomes in any given period—the domestic regulator will choose not to cheat by moving from point S toward point C.

In principle, an analogous repeated-interactions argument could be applied to a bargain to fully coordinate policy settings at point E. After all, if the domestic regulator cheats after agreeing to coordinate at point E, it experiences a one-time gain in welfare at point C. But if the bargain entails the foreign supervisor reverting to $\pi_N^* = 0$ for all remaining periods, then all future gains from operating at point E instead of point N will evaporate. Point E can be maintained if the domestic regulator cares enough about these future benefits from full coordination. The one-time benefit, relative to the bargaining costs to the nations for a bargain that maintains point S over time, could nonetheless turn out to be greater than the net benefit realized from achieving point C. Hence, maintaining a bargain to engage in full regulatory coordination could be more difficult than perpetuating alternative bargains that are less efficient but nonetheless preferable to no bargain whatsoever. There is indeed a range of potential bargains that the two supervisors could, in theory, adopt that would make them better off than allowing a noncoordinated regulatory race to the bottom.

As Gaspar and Schinasi note, however, "game theory has low predictive power," in the sense that "any socially feasible allocation that is individually rational can be obtained as an equilibrium." Where does this "low-predictive-power," dynamic-game-theoretic analysis leave the trilemma argument and, more generally, the feasibility of obtaining gains from international financial regulatory policy coordination? Regarding the trilemma issue, taking into account bargaining possibilities holds out more hope for dodging the unequivocal trilemma result. Even if transaction costs limit the scope for striking a bargain to engage fully in policy coordination that yields the most efficient possible outcome, more limited bargains that yield gains to participating regulators may still be feasible.

Concerning the more general gains-from-coordination issue, extending the static, single-shot-game-theoretic analysis utilized in Fig. 10.2 to a repeated-game setting continues to yield ambiguity. Allowing for repeated interactions adds to the standard set of requirements for coordination to pay off that banking and finance scholars have rediscovered, such as homogeneous policymaker preferences, a relatively small number of participants in the coordination agreement, monitoring and enforcement, and similar economic structures. Along with the importance of

policymaker patience, Gaspar and Schinasi add the applicable amount of time (number of repeated periods), degree of foresight and predictability of evolution of the policy environment, and transaction and bargaining costs. Accounting for dynamic-game-theoretic aspects of the financial policy coordination problem confronted by national regulators does not necessarily help or harm the case for coordination.

References

Acharya, Viral. 2003. Is the international convergence of capital adequacy regulation desirable? *Journal of Finance* 53: 2745–2781.

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.

Acharya, Viral, and Tanju Yorulmazer. 2008. Information contagion and bank herding. *Journal of Money, Credit, and Banking* 50: 215–231.

Acharya, Viral, Lasse Pedersen, Thomas Phillippon, and Matthew Richardson. 2010. Measuring systemic risk. Manuscript, Stern School of Business, New York University.

Acharya, Viral, Lasse Pedersen, Thomas Phillippon, and Matthew Richardson. 2011. Measuring systemic risk. In *Regulating Wall Street: The Dodd-Frank Act and the New Architecture of Global Finance*, ed. Viral Acharya and Matthew Ricardson, 87–119. New York: New York University.

Adrian, Tobias, and Markus Brunnermeier. 2009. CoVar, Federal Reserve Bank of New York Staff Report N. 348, Revised August.

Agur, Itai. 2014. Bank risk within and across equilibria. *Journal of Banking and Finance* 48: 322–333.

Aikman, David, Piergiorgio Alessandri, Bruno Eklund, Prasanna Gai, Sujit Kapadia, Elizabeth Martin, Nada Mora, Gabriel Sterne, and Matthew Willison. 2009. Funding liquidity risk in a quantitative model of systemic stability. Bank of England Working Paper No. 372, June 2009.

Albertazzi, Ugo, and Leonardo Gambacorta. 2009. Bank profitability and the business cycle. *Journal of Financial Stability* 5: 393–409.

Bae, Kee-Hong, G. Andrew Karolyi, and René Stulz. 2003. A new approach to measuring financial contagion. *Review of Financial Stdies* 16: 717–763.

Balasubramanyan, Lakshmi, and David VanHoose. 2013. Bank balance sheet dynamics under a regulatory liquidity ratio constraint. *Journal of Macroeconomics* 37: 53–67.

Barrell, Ray, E. Phillip Davis, Dilruba Karim, and Iana Liadze. 2010. Bank regulation, property prices, and early warning systems for banking crises in OECD countries. *Journal of Banking and Finance* 34: 2255–2264.

Barth, James, Gerard Caprio, and Ross Levine. 2014. Misdiagnosis: Incomplete cures for financial regulatory failures. Milken Institute Center for Financial Markets, October.

Beck, Thorsten, Radomir Todorov, and Wolf Wagner. 2013. Supervising cross-border banks: Theory, evidence, and policy. *Economic Policy* 28: 5–44.

Bernanke, Ben. 2008. Reducing systemic risk, Speech, Federal Reserve Bank of Kansas City Jackson Hole Conference, August 22.

Berrospide, Jose, Ricardo Correa, Linda Goldberg, and Friederike Niepmann. 2016. International banking and cross-border effects of regulation: Lessons from the United States. National Bureau of Economic Research Working Paper No. 22645, September.

Billio, Monica, Mila Getmansky, Andrew Lo, and Loriana Pelizzon. 2010. Measuring systemic risk in the finance and insurance sectors. MIT.

Bliss, Robert, and George Kaufman. 2006. Derivatives and systemic risk: Netting, collateral, and closeout. *Journal of Financial Stability* 2: 55–70.
Borio, Claudio. 2003. Towards a macroprudential framework for financial supervision and regulation? BIS Working Paper No. 128, Bank for International Settlements, February.
Borio, Claudio. 2009. The macroprudential approach to regulation and supervision. Vox. http://voxeu.org, April 14.
Born, Benjamin, Michael Ehrmann, and Marcel Fratzscher. 2010. Macroprudential policy and central bank communication. Center for Economic Policy Research Paper No. DP8094.
Boyer, Pierre, and Hubert Kempf. 2016. Regulatory arbitrage and efficiency of banking regulation. CESifo Working Paper No. 5878, April.
Brownlees, Christian, and Robert Engle. 2010. Volatility, correlation, and tails for systemic risk measurement. Working Paper, New York University, July.
Brunnermeier, Markus, Andrew Crockett, Charles Goodhart, Avinash Persaud, and Hyun Shin. 2009. The fundamental principles of financial regulation. Geneva Reports on the World Economy, #11, London: Center for Economic Policy Research.
Caballero, Ricardo, and Alp Simsek. 2013. Fire sales in a model of complexity. *Journal of Finance* 68: pp. 2549–2587.
Carlson, Mark, Burcu Duygan-Bump, and William Nelson. 2015. Why do we need both liquidity regulations and a lender of last resort? A persepective from Federal Reserve lending during the 2007–2009 financial crisis. Bank for International Settlements Working Paper No. 493, February.
Castrén, Olli, and Ilja Kristian Kavonius. 2009. Balance sheet interlinkages and macro-financial risk analysis in the euro area. European Central Bank Working Paper No. 1124, December.
Cetina, Jill, and Katherine Gleason. 2015. The difficult business of measuring banks' liquidity: Understanding the liquidity coverage ratio. Office of Financial Research Working Paper No. 15-20, U.S. Department of the Treasury, October 7.
Chakravorti, Sujit. 2000. Analysis of systemic risk in multilateral net settlement systems. *Journal of International Financial Markets, Institutions, and Money* 10: 9–30.
Čihák, Martin, and Klaus Schaeck. 2010. How well do aggregate prudential ratios identify banking system problems? *Journal of Financial Stability* 6: 130–144.
Clement, Piet. 2010. The term 'macroprudential': Origins and evolution. *BIS Quarterly Review, Bank for International Settlements*, March, 59–67.
Coase, Ronald. 1960. The problem of social cost. *Journal of Law and Economics* 3: 1–44.
Cocco, João, Francisco Gomes, and Nuno Martins. 2009. Lending relationships in the interbank market. *Journal of Financial Intermediation* 18: 24–48.
D'Hulster, Katia. 2011. Cross-border banking supervision: Incentive conflicts in supervisory information sharing between home and host supervisors. *Journal of Banking Regulation* 13: 300–319.
Dalen, Dag Morten, and Trond Olsen. 2003. Regulatory competition and multi-national banking. CESifo Working Paper No. 971, June.
Daníelsson, Jón. 2008. Blame the models. *Journal of Financial Stability* 4: 321–328.
Daníelsson, Jón, Kevin James, Marcela Valenzuela, and Ilknur Zer. 2016. Can we prove a bank guilty of creating systemic risk? A minority report. *Journal of Money, Credit, and Banking* 48: 795–812.
Davis, E. Phillip, and Dilruba Karim. 2008. Comparing early warning systems for banking crises. *Journal of Financial Stability* 4: 89–120.
Davis, E. Phillip, and Dilruba Karim. 2009. Macroprudential regulation—The missing policy pillar, Euroframe Conference on Economic Policy Issues in the European Monetary Union, June 12.
De Jonghe, Olivier. 2010. Back to basics in banking? A micro-analysis of banking system stability. *Journal of Financial Intermediation* 19: 387–417.
De Nicolo, Gianni, and Myron Kwast. 2002. Systemic risk and financial consolidation: Are they related? *Journal of Banking and Finance* 26: 861–880.

Dell'Ariccia, Giovanni, and Robert Marquez. 2006. Competition among regulators and credit market integration. *Journal of Financial Economics* 79: 401–430.

Devereux, Michael. 2014. New bank taxes: Why and what will be the effects. In *Taxation and Regulation of the Financial Sector*, ed. Ruud de Mooij and Gaëtan Nicodème, 25–54. Cambridge, MA: MIT Press.

Dia, Enzo, and David VanHoose. 2017. Bank taxes and loan Monitoring: A cautionary tale. Manchester School.

Diamond, Douglas, and Phillip Dybvig. 1983. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy* 91: 401–419.

Distinguin, Isabelle, Caroline Roulet, and Amine Tarazi. 2013. Bank regulatory capital and liquidity: Evidence from U.S. and European publicly traded banks. *Journal of Banking and Finance* 37: 3295–3317.

Eisenberg, Larry, and Thomas Noe. 2001. Systemic risk in financial systems. *Management Science* 47: 236–249.

Eldridge, Damien, Heajin Ryoo, and Axel Wieneke. 2015. Bank capital regulation with asymmetric countries. *Economic Record* 91: 79–90.

Elsinger, Helmut, Alfred Lehar, and Martin Summer. 2006a. Systemically important banks: An analysis for the European banking system. *International Economics and Economic Policy* 3: 73–89.

Elsinger, Helmut, Alfred Lehar, and Martin Summer. 2006b. Risk assessment for banking systems. *Management Science* 52: 1301–1314.

Engineer, Merwan, Paul Schure, and Mark Gillis. 2013. A positive analysis of deposit insurance provision: Regulatory competition among European Union countries. *Journal of Financial Stability* 9: 530–544.

Evans, Owen, Alfredo Leone, Mahinder Gill, Paul Hilbers, Winfrid Blaschke, Russell Krueger, Marina Moretti, Jun Nagayasu, Mark O'Brien, Joy ten Berge, and DeLisle Worrell. 2000. Macroprudential indicators of financial system soundness. Occasional Paper No. 192, International Monetary Fund, April.

Fecht, Falko, and Hans Peter Grüner. 2007. Financial integration and systemic risk. Unpublished Working Paper, Deutsche Bundesbank and University of Mannheim, January 22.

Folkerts-Landau, David. 1990. The case for international coordination of financial policy. In *International Policy Coordination and Exchange Rate Fluctuations*, ed. William Branson, Jacob Frenkel, and Morris Goldsten, 279–306. Chicago, IL: University of Chicago Press.

Freixas, Xavier, Bruno Parigi, and Jean-Charles Rochet. 2000. Systemic risk, interbank relations, and liquidity provision by the central bank. *Journal of Money, Credit, and Banking* 32: 611–638.

Furfine, Craig. 2003. Interbank exposures: Quantifying the risk of contagion. *Journal of Money, Credit, and Banking* 35: 111–128.

Gaspar, Vitor, and Garry Schinasi. 2010. Financial stability and policy cooperation. Occasional paper 1/2010, Bank of Portugal, July.

Gauthier, Céline, Alfred Lehar, and Moez Mouissi. 2010. Macroprudential capital requirements and systemic risk. Manuscript, Bank of Canada and University Calgary, July.

Hanson, Samuel, Anil Kashyap, and Jeremy Stein. 2011. A macroprudential approach to financial regulation. *Journal of Economic Perspectives* 25: 3–28.

Holthausen, Cornelia, and Thomas Rønde. 2004. Cooperation in international banking supervision. European Central Bank Working Paper No. 316, March.

Huang, Xin, Hao Zhou, and Haibin Zhu. 2009. A framework for assessing the systemic risk of major financial institutions. *Journal of Banking and Finance* 33: 2036–3049.

Huang, Xin, Hao Zhou, and Haibin Zhu. 2010. Systemic risk contributions. Working Paper, University of Oklahoma, Board of Governors of the Federal Reserve System, and Bank for International Settlements, August.

Iori, Giulia, Saqib Jafarey, and Francisco Padilla. 2006. Systemic risk on the interbank market. *Journal of Economic Behavior and Organization* 61: 525–542.

Jacob, Punnoose, and Anella Munro. 2016. A macroprudential stable funding requirement and monetary policy in a small open economy. Crawford School of Public Policy Centre for Applied Macroeconomic Analysis Working Paper 23/2016, May.

Jorion, Philippe, and Gaiyan Zhang. 2009. Credit contagion from counterparty risk. *Journal of Finance* 64: 2053–2087.

Kane, Edward. 2010a. Missing elements in U.S. financial reform: A Kübler-Ross interpretation of the inadequacy of the Dodd-Frank Act. Working Paper, Boston College, August.

Kane, Edward. 2010b. Redefining and containing systemic risk. Atlantic Economic Journal 38: 251–264.

Kashyap, Anil, Richard Berner, and Charles Goodhart. 2010. The Macroprudential Toolkit. Chicago Booth Initiative on Global Markets Working Paper No. 60, December.

Kopecky, Kenneth, and David VanHoose. 2012. Can capital requirements induce private monitoring that is socially optimal? *Journal of Financial Stability* 8: 252–262.

Kretzschmar, Gavin, Alexander McNeil, and Axel Kirchner. 2010. Integrated models of capital adequacy—Why banks are undercapitalized. *Journal of Banking and Finance* 34: 2838–2850.

Lehar, Alfred. 2005. Measuring systemic risk: A risk management perspective. *Journal of Banking and Finance* 29: 2577–2603.

Liebmann, Eva, and Joe Peek. 2015. Global standards for liquidity regulation. Current Policy Perspectives, Federal Reserve Bank of Boston, July.

Morrison, Alan, and Lucy White. 2009. Level playing fields in international financial regulation. *Journal of Finance* 64: 1099–1142.

Persaud, Avinash. 2009. Macro-prudential regulation, Crisis Response #6, World Bank, July.

Rochet, Jean-Charles. 2010. Systemic risk: Changing the regulatory perspective. *International Journal of Central Banking* 6: 259–276.

Rochet, Jean-Charles, and Jean Tirole. 2006. Two-sided markets: A progress report. *Rand Journal of Economics* 37: 645–667.

Rochet, Jean-Charles, and Jean Tirole. 2008. Competition policy in two-sided markets, with special emphasis on payment cards. In *Handbook of Antitrust Economics*, ed. Paulo Buccirossi, 543–582. Cambridge, MA: MIT Press.

Schoenmaker, Dirk. 2011. The financial trilemma. *Economics Letters* 111: 57–59.

Segoviano, Miguel, and Charles Goodhart. 2009. Banking stability measures. International Monetary Fund Working Paper WP/09/04, January.

Sin, Hans-Werner. 2001. Risk taking, limited liability, and the competition of bank regulators. NBER Working Paper 8669, December.

Tarashev, Nikola, Claudio Borio, and Kostas Tsatsaronis. 2010. Attributing systemic risk to individual institutions. Bank for International Settlements Working Paper No. 308, May.

Temesvary, Judit. 2015. The role of regulatory arbitrage in U.S. banks' international lending flows: Bank-level evidence. FIW Working Paper No. 151, March.

VanHoose, David. 2011a. Systemic risk and macroprudential bank regulation: A critical appraisal. *Journal of Financial Transformation* 33: 45–60.

VanHoose, David. 2011b. Regulation of bank management compensation. In *Financial Market Regulation: Legislation and Implications*, ed. John Tatom, 163–183. New York: Springer.

VanHoose, David. 2011c. Regulatory constraints on performance-based managerial compensation, bank monitoring, and aggregate loan quality. *Atlantic Economic Journal* 39: 315–328.

VanHoose, David. 2013. Evaluating the centralized-layers approach to U.S. federal financial regulation. Mercatus Center, George Mason University, July 15.

VanHoose, David, and Gordon Sellon, Jr. 1989. Daylight overdrafts, payments system risk, and public policy. *Federal Reserve Bank of Kansas City Economic Review*, September/October.

Veljanovski, Cento. 2010. Economic approaches to regulation. In *The Oxford Handbook of Regulation*, ed. Robert Baldwin, Martin Cave, and Martin Lodge. Oxford: Oxford University Press.

von Peter, Goetz. 2009. Asset prices and banking distress: A macroeconomic approach. *Journal of Financial Stability* 5: 298–319.

Wagner, Wolf. 2008. The homogenization of the financial system and financial crises. *Journal of Financial Intermediation* 17: 330–356.
Wagner, Wolf. 2009. In the quest of systemic externalities: A review of the literature. Working Paper, Tilburg University, February.
Wagner, Wolf. 2010. Diversification at financial institutions and systemic crises. *Journal of Financial Intermediation* 19: 373–386.
Weinberg, John. 2002. Competition among bank regulators. *Federal Reserve Bank of Richmond Quarterly* 88 (Fall): 19–36.
White, Lawrence. 1994. On the international harmonization of bank regulation. *Oxford Review of Economic Policy* 10: 94–105.
White, William. 2008. Past financial crises, the current financial turmoil, and the need for a new macrofinancial stability framework. *Journal of Financial Stability* 4: 307–312.
Wyplosz, Charles. 2009. The ICMB-CEPR Geneva Report: 'The Future of Financial Regulation,' Vox, January 27 http://www.voxeu.org/index.php?q=node/2872
Zhou, Chen. 2010. Are banks too big to fail? Measuring systemic importance of financial institutions. *International Journal of Central Banking* 6: 205–250.