ESSAYS ON BANKING AND FINANCIAL REGULATIONS

by

Hai Xuan Nguyen

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

July, 2015

© Hai Xuan Nguyen 2015

All rights reserved
Abstract

This dissertation consists of three essays on banking and financial regulations. Using analytical frameworks and historical perspectives, these essays investigate some of the most pressing issues underlying the regulatory structure of the United States, including regulatory competition, “too big to fail,” and shadow banking. The results provided herein help foster the discussion on future financial reforms.

The first essay analyzes the welfare impact of supervisory shopping in the banking sector. Supervisory shopping leads to a welfare-increasing “race to the top” among supervisors, if strong supervision increases banks’ access to deposits by signaling that banks have a less risky balance sheet. However, if deposits are subsidized through an underpriced financial safety net, banks shop for a weak supervisor to maximize gains from risk-shifting. Supervisory shopping then precipitates a “race to the bottom.” In an extension of the framework, contingent convertible bonds (CoCos) prevent banks from risk-shifting, and supervisory shopping becomes socially desirable again.

The second essay, which is a joint effort with Chang Ma, examines the optimal prudential regulations of large banks in the presence of an underpriced bailout guarantee. We evaluate policies that aim to mitigate the social costs of a bailout, including a capital requirement, direct cap on size, and tax on size. We find that each policy proposal improves social welfare in comparison with its absence; however, it cannot achieve the first-best level. In particular, bank size regulation inhibits banks’ scale economies, because it curbs bailout costs by
limiting bank size. To eliminate the use of public funds and the associated costs, we consider the issuance of CoCos to absorb banks' losses in the bad state. We find that a socially optimal investment scale can be implemented when CoCos are required.

The third essay, which is a joint effort with Dang Du, reviews the history of financial regulatory developments in the U.S. via an examination of major regulatory reforms. While drawing on historical narratives of financial overhauls, we pay special attention to the issue of regulatory leakages that have led to the rise—in various forms—of unregulated short-term funding markets, ineffective regulatory regimes, and shadow banking. Our discussion highlights the fact that these leakages, which have been regarded as leading causes of the most recent financial crisis (2007–09), have deep roots in U.S. history.

**Keywords:** Supervisory Shopping, Banking Supervision, Regulatory Competition, Race to the Top, Race to the Bottom

Too Big To Fail, Bailout, CoCos, Size Regulation


**JEL classification:** G01, G18, G21, G28, F36, L51, N21, N22

**Primary Reader:** Olivier Jeanne

**Secondary Reader:** Laurence M. Ball
Acknowledgement

My journey towards the completion of the Ph.D. program has been rewarding, thanks to countless people. Without them and their help and support, I would never have been able to finish this dissertation.

I would like to express my deepest gratitude to my advisor, Olivier Jeanne. His wisdom, direction, and tolerance afforded me the opportunity to discover my own path. When my steps faltered, he gave me the guidance to recover and stay on track. With his insightful questions and constructive criticisms, he taught me how to conduct research, solidify my progress, and formalize my results. His comments and suggestions on each of the many revisions of this manuscript were always trenchant and instructive—and more detailed than I could ever have hoped for. I count myself among the most fortunate students to have benefited so tremendously from his patience, encouragement, and supervision.

I am also grateful to Laurence Ball, Jon Faust, Anton Korinek, Hülya Eraslan, and Joseph Harrington for reading numerous versions of my research, participating in my presentations, commenting in detail on my strengths and weaknesses, and helping me cultivate and clarify my ideas.

I am thankful for Dang Du and Chang Ma for their collaborations. With his specialized training in public policy, Dang possesses a great deal of institutional knowledge regarding the U.S. financial system, from which I have greatly benefited. I am grateful to Chang for the long discussions that helped me sort and verify the technical details of this manuscript. It has been a pleasure and a privilege working with both of them.

I would also like to thank all of my friends at Johns Hopkins who have
helped me through graduate school. Special thanks go to Hou Wang, whose friendship has enriched my life in more ways than I can name. I would also like to thank Natsuki Arai for his advice and support during the job market process. I am grateful to the members of the Macro-Finance Discussion Group and the participants of the Macro seminars for their feedback on my work.

I am especially thankful to my two childhood friends, Dat Le and Huy Nguyen. While each of us has followed a different path in life, I have always been able to count on their unconditional friendship to get me through the toughest times.

Finally, I could not be who I am or where I am if it were not for the love and support of my family. My father, Hoan Nguyen, is my guiding light. My mother, Van Phung, is the bedrock of all of my accomplishments, big and small, in my personal life and in my professional endeavors. My younger brother, Hiep Nguyen, is the fuel to my perseverance. My wife, Phuong Tran, is the source of my inspiration. And my son, Dang Nguyen, is the joy of my life. To them, this dissertation is dedicated.
# Table of Contents

**Front Matter**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgement</td>
<td>iv</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>viii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td>List of Figures</td>
<td>x</td>
</tr>
</tbody>
</table>

1 Supervisory Shopping in the Banking Sector: When is it Socially Desirable?  

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Supervisory Shopping in Practice</td>
<td>5</td>
</tr>
<tr>
<td>1.2.1 The U.S. Experience</td>
<td>6</td>
</tr>
<tr>
<td>1.2.2 The International Experience</td>
<td>9</td>
</tr>
<tr>
<td>1.3 Related Literature</td>
<td>11</td>
</tr>
<tr>
<td>1.4 The Model</td>
<td>14</td>
</tr>
<tr>
<td>1.5 Laissez-Faire Debt Contract</td>
<td>18</td>
</tr>
<tr>
<td>1.6 Regulation and Supervision of Banks</td>
<td>20</td>
</tr>
<tr>
<td>1.6.1 Supervision Without An Ex-post Bailout</td>
<td>20</td>
</tr>
<tr>
<td>1.6.2 Capital Requirement and Bailout Guarantee</td>
<td>31</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>1.7</td>
<td>Extensions</td>
</tr>
<tr>
<td>1.7.1</td>
<td>CoCos and The Best of Both Worlds</td>
</tr>
<tr>
<td>1.7.2</td>
<td>Shopping in An International Context</td>
</tr>
<tr>
<td>1.8</td>
<td>Conclusion</td>
</tr>
<tr>
<td>2</td>
<td>Too Big To Fail: Toward an Optimal Regulation</td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>2.2</td>
<td>Literature Review</td>
</tr>
<tr>
<td>2.3</td>
<td>The Model</td>
</tr>
<tr>
<td>2.4</td>
<td>Moral Hazard of Government Bailouts</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Too Big To Fail in the Absence of Regulation</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Capital Requirement</td>
</tr>
<tr>
<td>2.5</td>
<td>Bank Size Regulation: Better, But Not the Best</td>
</tr>
<tr>
<td>2.5.1</td>
<td>A Cap On Size</td>
</tr>
<tr>
<td>2.5.2</td>
<td>A Tax On Size</td>
</tr>
<tr>
<td>2.6</td>
<td>CoCos: Toward Better Regulation of Large Banks</td>
</tr>
<tr>
<td>2.7</td>
<td>Extension: Investment Concentration</td>
</tr>
<tr>
<td>2.8</td>
<td>Conclusion</td>
</tr>
<tr>
<td>3</td>
<td>A Historical Walk Through Financial Reforms: Lessons for the Future</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>3.2</td>
<td>The National Banking Era, 1863–1913</td>
</tr>
<tr>
<td>3.2.1</td>
<td>A Historical Context</td>
</tr>
<tr>
<td>3.2.2</td>
<td>National Banking and Its Weaknesses</td>
</tr>
<tr>
<td>3.3</td>
<td>The Fed: History, Politics, and Leakages</td>
</tr>
</tbody>
</table>
3.3.1 A Prelude to the Federal Reserve Act . . . . . . . . . . . 108
3.3.2 The Fed’s Early Years, 1913–1933 . . . . . . . . . . . . 114
3.4 Financial Reforms in the 1930s . . . . . . . . . . . . . . . . . 121
3.5 Deregulation Amidst Regulatory Failures . . . . . . . . . . . . 126
3.6 The 2007–09 Crisis and the Re-Reforms . . . . . . . . . . . . 132
  3.6.1 The Rise of Shadow Banking . . . . . . . . . . . . . . . 132
  3.6.2 It’s a Liquidity Crisis, Again! . . . . . . . . . . . . . . . 135
  3.6.3 Regulatory Re-Reform: The Dodd–Frank Act of 2010 . . 139
3.7 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 144

Bibliography 146

A Appendix for Chapter 1 158
  A.1 Two Supervisors and A Yardstick Competition . . . . . . . . 158
  A.2 Proof of Proposition 1.7.1 . . . . . . . . . . . . . . . . . . . 159
  A.3 Proof of Proposition 1.7.3 . . . . . . . . . . . . . . . . . . . 161
  A.4 Motivation by Punishment? . . . . . . . . . . . . . . . . . . 162
    A.4.1 A Single Supervisor . . . . . . . . . . . . . . . . . . . . 164
    A.4.2 Two Supervisors And More . . . . . . . . . . . . . . . 168

B Appendix for Chapter 2 171
  B.1 Proof of Proposition 2.6.1 . . . . . . . . . . . . . . . . . . . 171

Curriculum Vitae 173
List of Tables

2.1 Banks’ payoffs and joint distribution of returns . . . . . . . . . . . . . . . . . . . . . . . . . . . 84
List of Figures

1.1 Projects’ Returns ........................................ 15
Chapter 1

Supervisory Shopping in the Banking Sector: When is it Socially Desirable?

1.1 Introduction

Whenever a banking or financial crisis occurs, the effectiveness of financial regulation and supervision is called into question. For instance, immediately after the most recent global financial crisis that erupted in 2007, many experts and policymakers urgently called for a wholesale change to the regulatory and supervisory apparatus, which subsequently resulted in a major redesign of the regulatory and supervisory environment in the United States as well as in the Eurozone and other financial centers. The optimal design, however, still very much remains an elusive concept.

Supervisory shopping, among others, poses a great challenge to crafting an optimal design in today’s increasingly integrated world. Financial institutions, which typically have access to multiple supervisors, can and often do switch supervisors at a relatively low cost. This opens the door to supervisory
shopping where financial institutions pick and choose their “favorite” supervisors. Researchers and policymakers alike have acknowledged the presence of supervisory shopping in practice. Many of them—for instance, White (2009) and Johnson and Kwak (2011)—argue that it leads to unhealthy competition among supervisors and reduces social welfare. Others have come to its defense and, for instance, argue that financial institutions’ ability to switch can act as a healthy constraint on the authorities’ policy-making process (Greenspan, 1998; Wilcox, 2006). The jury, one might say, is still out on the impact of supervisory shopping in the banking sector.

In this paper, I propose a three-tier principal-agent framework of banking supervision to study the welfare impacts of supervisory shopping. The first tier consists of banks that borrow from depositors and make investment decisions. The second tier consists of supervisors who examine, identify, and resolve risky banks. To accomplish her tasks, each supervisor must acquire skills and expertise as well as exert effort—all variables that are often not verifiable in the court of law. When there are multiple supervisors, I assume that each bank is subject to examination by only one supervisor of its choice. Finally, the third tier consists of a principal, Congress, who offers supervisors incentive contracts and imposes explicit regulations on banks to maximize social welfare. Regulation in this paper takes the form of a minimum capital requirement, which can prevents bankers’ excessive risk-taking. I then examine the equilibrium of the framework, depending on the availability of a guaranteed bailout for banks.

A crucial element of the framework is that an incentive contract for supervisors is contingent on the numbers of supervised and resolved banks but not on supervisory effort. This is because supervisory effort cannot be verified in
court. This captures a reality that market participants can get a sense of how effective a supervisor is but cannot easily describe the basis for this assessment in a way that is provable in a court of law. By contrast, the numbers of supervised and resolved banks are both observable and verifiable. The supervisory framework can be designed so as to use the verifiable information as a proxy for the supervisory effort.

The main result of the framework is that supervisory shopping has a positive impact on social welfare when Congress can credibly commit to a no-bailout policy. The intuition is that without a bailout, bankers require a strong supervisor who can catalyze market confidence in the banking system. That, in turn, allows bankers to increase leverage and raise return on equity. Capitalizing on bankers’ preference, Congress then incentivizes supervisors by letting bankers “vote with their feet” and rewarding supervisors proportionally to the number of banks they receive. Supervisors, as a consequence, enter a “race to the top” in effort levels to attract banks. However, bank failures are socially costly. An ex post welfare-maximizing Congress that finds it optimal to rescue insolvent banks thus instills a full ex ante expectation of a bailout. That distorts bankers’ preference toward weaker supervision, which allows bankers to maximize gains from risk-shifting. Supervisory shopping, in this case, precipitates a “race to the bottom” among the supervisors.

My first extension of the framework is to allow bankers to issue contingent convertible bonds (CoCos). They are hybrids of equity capital and debt that can absorb losses in a crisis. When bankers issue CoCos, bankruptcy and the associated social cost are avoided without government interventions, thus, making the no-bailout commitment credible. Accordingly, supervisory shopping
helps to improve social welfare. The intuition is straightforward: CoCo-issuing bankers act as if they were issuing deposits (debts) and therefore prefer high-effort supervisors. If their investments fail, CoCos automatically convert into equity, eliminating the need for bankruptcy and thus the \textit{ex ante} anticipation of a bailout.

My second extension of the framework considers supervisory shopping in an international context. The extended framework includes two countries—one big and one small—each with its own set of Congress, supervisors, and depositors. For analytical convenience, I assume all bankers are shareholders in the big country where they raise deposits and invest. Bankers, however, can choose to locate their headquarters in either country, where they will be subject to the regulation and supervision of that country on a consolidated basis. The small country receives a share $\alpha$ of profits of banks that register there and helps pay for a fraction $\beta$ of the costs of bailing out these banks when applicable. The analysis then reveals two noteworthy results. First, when both countries commit to not bail out banks, supervisory efforts may become inefficiently high. That is because bankers prefer to charter in a country with strong supervision so that they can save on equity and have a high leverage ratio. The two countries then compete to attract bankers by raising supervisory efforts, albeit to socially inefficient levels. In comparison with single-country cases, oversupervision occurs here because of the existence of two competing supervisory regimes. This, therefore, speaks in favor of international coordination of supervision. Second, when bailouts are available, the relative magnitudes of $\alpha$ and $\beta$ can affect the two countries’ regulatory requirements. In particular, for a large $\alpha$ and small $\beta$, the small country has a strong incentive to attract banks, even if they are
risky. That forces the two countries into a “race to the bottom” in regulatory requirements.

My results are relevant in a wide range of circumstances. In the US, for instance, supervisory competition has long been a part of the country’s financial landscape, tracing back to 1864 when the system of dual banking was first created. Any commercial bank in the US has been able to select a supervisor of choice via its chartering decision. As another example, banks and financial institutions in the highly integrated Eurozone market, where the formation of a banking union is forthcoming, will naturally consider the variations among national authorities and between European supervisors inside and outside of the Eurozone. Consequently, supervisory competition as described in this paper requires careful consideration.

The rest of the paper proceeds as follows. Section 1.2 describes supervisory shopping in practice in the US and in an international context. Section 1.3 discusses the related literature. Section 1.4 sets up the framework. Section 1.5 provides the laissez-faire benchmark. Section 1.6 analyzes the regulation and supervision of banks. Section 1.7 extends the framework to study the use of CoCos and supervisory shopping across borders. Finally, Section 1.8 concludes the paper.

1.2 Supervisory Shopping in Practice

This section highlights the importance of supervisory shopping by examining the US and the Eurozone banking environments. In the US, supervisory shopping has long been a part of the financial landscape, arguably since the passage of the
National Banking Act of 1864 that created a dual system of state and federal commercial banks. In the Eurozone, a banking union is being formed. While regulatory standards there are to be aligned, supervisory practices continue to vary across member countries. Supervisory shopping, therefore, will become an important issue in the area.

The discussion below will address three forms of supervisory shopping in the US and in an international context and highlight their connection with the model presented in Section 1.4. The three forms include: (1) an institution in a well-defined sector (e.g. commercial banking) chooses among the supervisors who oversee that sector; (2) a conglomerate that is active in multiple lines of business (i.e. banking and insurance) and is supervised by different agencies may select one to be its primary supervisor; and (3) an international institution may select a national authority that most fits its need by moving its headquarters across countries. While the section on “The U.S. Experience” will showcase the first two forms, the section on “The International Experience” will do the last.

1.2.1 The U.S. Experience

In the US, banks and financial institutions have always retained the option to “shop” for regulators and supervisors of their choice since the end of the Civil War. Not only is this option frequently exercised, but also it can manifest itself in different forms.

The first and most straightforward form of supervisory shopping in the US is switching charters by commercial banks, which can leave detrimental effects on the budgets of supervisory agencies. Between 1950 and 1977, there were 1,828 switches from state to federal agencies or from one federal agency to another.
Between 1977 and 2003, more than 10% of all banks did this (Rosen, 2005). When an agency loses or gains a bank, its annual budget may fluctuate to a great extent. For instance, California Superintendent of Banks lost 30% of revenues in 1968 when Wells Fargo Bank converted to a national bank (Scott, 1977). Similar impacts occurred when Chase Manhattan Bank converted from a New York charter to a national one in 1965 and when Wachovia Bank in North Carolina became a national bank in 1968 (Scott, 1977). Another example includes the Office of the Comptroller of the Currency (OCC)—a federal regulator/supervisor—which lost 2% of its budget when Chase Manhattan Bank merged with Chemical Bank and decided to have New York’s state charter in 1995 (Rosen, 2003). Also, the State of New York Banking Department lost 30% of its annual budget when J.P. Morgan Chase Bank and HSBC Bank USA decided to convert from the New York charters to national charters as recently as 2004 (Blair and Kushmeider, 2006).

Because the budgets of state and federal agencies often depend on the number and the sizes of banks they supervise, they frequently adjust their regulatory and supervisory standards to preserve and increase membership (Scott, 1977; White, 1982; Calomiris, 2000, and others). Before World War I, for instance, state and federal banking authorities repeatedly lowered barriers to entry and reduced capital requirements (White, 1982). Supervision was also believed to be laxer at state levels than at the federal level (White, 2009). Although numerous reforms have made banking regulations more uniform across state and federal levels since the Great Depression, Scott (1977) notes that agencies continue to compete to promote their banks’ profitability via discretionary actions, such as “the exercise of approval powers over mergers or branches, and questions
of statutory interpretation within the bounds allowed by language and legal conscience.”

The second form of supervisory shopping is financial conglomerates’ strategic selection of one primary federal supervisor. For instance, before the passage of the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 (Dodd-Frank), a given conglomerate could easily become either a bank/financial holding company, which was supervised by the Federal Reserve (Fed), or a thrift holding company, which was supervised by the Office of Thrift Supervision (OTS). The choice of the primary supervisor can be essential to the profitability of a financial institution because the primary supervisor is responsible for evaluating and determining the company’s overall risk profile. By selecting the “right” primary supervisor, for instance, a conglomerate may be able to achieve its desired level of net exposure across all of its holdings.

Conspicuous examples of this shopping behavior involve the OTS and many of its supervised entities. According to the Government Accountability Office (2007), the OTS was the primary supervisor for a substantial number of “large” and “complex” financial conglomerates that had “primary businesses other than those traditionally engaged in by thrifts.” Those include Countrywide, IndyMac, and WaMu, which all converted to thrift holding companies before the 2007–09 crisis and then failed due to excessive risk-taking under the OTS’s watch. The most notorious one was the global insurance giant AIG, whose failure nearly brought the entire system to the brink of collapse (Johnson and Kwak, 2011; Sorkin, 2010). With $259 billion in consolidated assets, AIG became a thrift holding company in 1999 by forming a $22 million-dollar thrift (Office of Thrift
Supervision, 1999). The crisis later made clear that the OTS had inadequate expertise to supervise AIG’s vast and complex business activities, which included a wide range of financial and insurance-related products (Johnson and Kwak, 2011). OTS’s lack of expertise, nonetheless, might have been the very reason AIG and others became thrift holding companies.

1.2.2 The International Experience

International financial integration has been a clear trend in recent decades, especially among the European Union (EU) countries (Lane and Milesi-Ferretti, 2003; Dermine, 2006). For this reason, supervisory shopping across borders is and will continue to be a pertinent concern.

Supervisory shopping across borders occurs when financial institutions register their headquarters strategically so as to be supervised by a national authority of choice. A notorious example is the Bank of Credit and Commerce International (BCCI), which operated the majority of its businesses from London but was headquartered in Luxembourg. At the time, Luxembourg did not have any law that subjects BCCI to supervision on a consolidated basis. According to Corrigan, Mattingly, and Taylor (1992), BCCI’s particular structure allowed it to elude “normal banking oversight” and engage in fraudulent activities. BCCI was eventually forced to close and liquidate in 1991, but the liquidation process lasted for 21 years and costed more than $1.7 billion (Croft, 2012, May 17).

Although in the wake of BCCI, national banking authorities have together developed a common approach toward consolidated supervision of international banks, supervisory shopping across borders was not sufficiently addressed. Under the aegis of the Basel Committee on Banking Supervision (1983, 1997),
many countries, including all of the members of the EU, have adopted the principle of home-country supervision as outlined in the 1983 "Basel Concordat" and the 1997 "Core Principles for Effective Banking Supervision." According to the principle, the national banking authority with whom an international bank registers its headquarters is in charge of supervising the bank on a consolidated basis. While this principle significantly reduces jurisdictional confusion, it concentrates the supervisory power into a single national authority. In addition, as there has yet to be comprehensive international supervisory standards, each country retains the freedom to continue its own approach to supervision. As a result, banks continue to have strong incentives to seek supervisors who are most accommodating to their businesses. They can do so simply by either relocating their headquarters or engaging in cross-border mergers and acquisitions. Holthausen and Rønde (2004) have also argued the same point in their paper.

Notably, supervisory integration in the recently created banking union in Europe, albeit a step in the right direction, continues to leave room for supervisory shopping because of a large degree of delegation to national authorities. Under the current regulatory framework, the European Central Bank (ECB) will uniquely provide supervision on a consolidated basis for all bank and financial holding companies that are established within the Eurozone (Union, 2013). However, because the ECB has limited resources, it will only directly supervise large and systemic institutions on a broad scale. National supervisory authorities will then continue to play a crucial role in banking supervision. For instance, not only are they responsible for most day-to-day supervisory activities, but also national authorities will be in charge of executing bank resolutions. Furthermore, the current proposal solely addresses the commercial
banking sector, leaving out the rest of the financial industry. It is unclear how other financial institutions, including securities firms and insurance providers, will be supervised under the new banking union.

The structure of a banking union in Europe will share many similarities with the U.S. structure in the past. For instance, a bank that operates in the integrated European market continues to have various options for supervision. It can be chartered by the ECB while locating in one of the Euro member countries; be chartered by a EU member country outside of the Eurozone and open branches anywhere in Europe; or own a bank within the Eurozone while forming banking/financial holding companies outside of the Eurozone. For this reason, some degree of supervisory shopping will persist in the new banking union. The model and subsequent analysis in this paper will accordingly help shed light on the costs and benefits of such competition.

1.3 Related Literature

Three relevant bodies of literature are regulatory economics, banking regulation and supervision, and finally, financial regulatory and supervisory competition.

In the regulatory economics literature, principal-agent frameworks are frequently used to capture and study the incentives of regulated entities [see Laffont and Martimort (2009) for a survey]. Beginning with the seminal work by Tirole (1986), the traditional two-tier regulator-regulated framework was extended to include a supervisory layer. Laffont and Tirole (1986, 1991) develop the idea further and emphasize the welfare loss when the single supervisor is self-interested. In this paper, I employ a similar approach to supervision by conceptualizing
the financial market as a three-tier system. Unlike previous research, however, I endogenize the supervisors’ effort and focus on their incentives when they have to compete with one another for income.

Using a three-tier principal-agent framework built on the aforementioned research, Laffont and Martimort (1999) are the first to consider a setting with multiple supervisors. In their paper, the supervised firm has two separate bits of private information, each of which is revealed to the supervisory authority with an exogenous probability. The authors find that welfare is improved with two independent supervisors in place. That is because each of the two supervisors, who do not collude with each other, has less information about the regulated firm and henceforth finds it harder to enter a side contract with the firm. Boyer and Ponce (2012) use the same modeling approach to study banking supervision and find similar results.

My research differs from their papers because bankers in my framework have the ability to “shop” for only one supervisor. This assumption is a key element that enables my framework to examine the impact of supervisory shopping in the banking sector. It aligns supervisors’ incentives with that of the supervised banks, which, in turn, leads supervisory competition toward bankers’ benefits. As supervisors strive to attract banks by catering to their demand, it is conceptually unclear whether such a behavior is socially desirable (Kane, 1984; Dermine, 1991; White, 1994). With a framework allowing for supervisory shopping, my paper helps to analytically determine the impact of this type of supervisory competition.

To bring the paper in line with the literature on banking regulation and supervision, I include a regulatory instrument in the form of a minimum capital
requirement. Following *Hellmann, Murdock, and Stiglitz* (2000), my framework features costly equity capital due to bankers’ high discount rate. I find that a sufficiently high capital requirement can limit bankers’ risk-taking behaviors. Similar results can be found in a number of other papers, including Rochet (1992), Blum (1999), Matutes and Vives (2000), John, Saunders, and Senbet (2000), and Repullo (2004). In addition, I show that supervision is complementary to regulation in achieving the optimal level of social welfare if the supervisory effort is verifiable. *Campbell, Chan, and Marino* (1992) have the same approach to regulation and supervision, and they accordingly reach similar conclusions.

Furthermore, minimum capital requirements are a typical instrument for studying regulatory competition. Examples include *Dell’Ariccia and Marquez* (2006) and *Sinn* (2001). While Dell’Ariccia and Marquez model banks with international loan portfolios, Sinn models banks that raise deposits both at home and abroad. In both of these papers, national authorities use capital requirements to mitigate banks’ risk-taking but fail to account for externalities of their policies in the foreign country. As a result, capital requirements are lower in equilibrium relative to what would have been set by an internationally centralized authority.

Departing from this trend, my paper shares a similarity with the work by *Acharya* (2003) when considering a common capital requirement for all banks within our respective framework. As he asks whether the convergence of capital requirements is desirable, Acharya examines regulatory competition beyond the use of this instrument. Instead, he focuses on national authorities’ closure policy, which is modeled as a probability of resolving an insolvent bank. His
paper finds that there is a “regression toward the worst forbearance” because each national authority races to protect only domestic banks. Similar results can also be found in Repullo (2001) and Holthausen and Rønde (2004).

My paper, however, focuses on supervision and examines the respective supervisory competition. Weinberg (2002) is the closest paper to mine in this aspect. In his work, Weinberg also considers banks’ ability to select a supervisor and the resulting supervisory competition. Different from my approach in which supervisors are self-interested agents, Weinberg describes supervisors as policymakers whose objective is to maximize the difference between banks’ profits and the supervisory costs. He finds that competition among supervisors leads to efficient outcomes if each supervisor is also a deposit insurance provider. Otherwise, competition will lead to a “race to the bottom.”

1.4 The Model

Let us consider a game-theoretic 2-period model of banking regulation and supervision with a continuum of identical depositors with mass 1 and a continuum of entrepreneurial bankers with mass $b$, both of whom are risk-neutral. Banking regulation and supervision, which will be discussed in detail later, are managed by a welfare-maximizing entity, referred to in this model as “Congress”.

I assume that bankers have access to investment opportunities that are not available to depositors. The natures of the investments, banking activities and bankers and depositors’ utilities are described in details below.

Investment Projects: There are two types of investment projects - “safe” and “risky,” both of which require at Date 1 an indivisible investment of 1. At
Date 2, a safe project yields $\varepsilon$ with probability $\bar{r}$; and $\varepsilon$ otherwise. Meanwhile, a risky project yields $R$ with probability $r$; and $\varepsilon$ otherwise. I assume that return shocks are perfectly correlated across all banks.

Assume that $R > \varepsilon > 1 > \underline{\varepsilon} > 0$; and $1 > \bar{r} > \underline{r} > 0$. Figure 1.1 depicts the projects' payoffs with respective probabilities.

**Banking Activities:** At Date 1, depositors receive an initial endowment $w$, which can be made available for bankers to borrow ($w > b$). Depositors have rational expectations and preferences linear in consumption:

$$U^d(c) = c_1 + c_2,$$

where $c = (c_1, c_2)$ represents depositors' consumptions at Date 1 and 2.

Also at Date 1, each banker receives 1 unit of endowment that can be directed toward investment. Given that bankers can raise additional funding from depositors, they may invest only a fraction of their endowment as equity capital.
and consume the rest during Date 1. Their utility is given as follow:

\[ U^b(c) = c_1 + \frac{1}{\delta} c_2, \]

where \( \delta - 1 \geq 0 \) represents bankers’ intertemporal discount rate. Alternatively, \( \delta \) can be interpreted as the internal cost of equity for bankers.

Given the investment opportunities and the cost of equity, bankers would like to borrow from depositors to invest. Let the debt contract be of the form \((k, i)\) such that each banker is to borrow \(1 - k\) from depositors at Date 1, and to repay \((1 - k)(1 + i)\) at Date 2. I assume that bankers make a take-it-or-leave-it offer, and hence, extract all the rent.

Moreover, the following assumptions are imposed to reflect the general nature of a lending market: (1) bankers have limited liability; (2) the type of bankers’ project is private information; and (3) in case of a default, all of their assets are liquidated and distributed to depositors. These assumptions imply, in particular, that bankers will not strategically default if they can meet their debt obligations in period 2.

I assume that bankruptcy is costly so that depositors can recover only \( \lambda < 1 \) units for every 1 unit of the seized assets. This assumption can be rationalized, for instance, by costly state verification as in Townsend (1979).

Let us denote:

\[ \Delta P = \bar{r}_\varepsilon - r R, \]

\[ \Delta r = \bar{r} - r. \]
I assume that:

\[ \bar{r} \varepsilon + (1 - \bar{r}) \lambda \varepsilon > \delta, \quad (1.1) \]

\[ 1 > r R + (1 - r) \xi, \quad (1.2) \]

and \( \Delta r > \Delta P. \quad (1.3) \)

Inequality (1.1) implies that the expected return of a safe project, even when accounting for the bankruptcy cost, is higher than bankers’ internal cost of equity. Thus, safe projects are socially desirable. On the other hand, inequality (1.2) implies that risky projects are not socially beneficial. This assumption helps reduce the number of cases to be considered, but does not affect the essence of the results as long as the return of a risky project is lower than that of a safe one.

Finally, conditional on 0 interest rate, the expected profit from risky projects is higher than from safe ones by (1.3).

**Congress:** Because projects are indivisible and require exactly 1 unit of investment, \( k \) can be interpreted as a capital ratio. In our model, Congress can influence banking activities by setting a minimum capital ratio \( \underline{k} \). The objective of Congress is to maximize social welfare, which is defined as a sum of bankers’ and depositors’ utilities:

\[ W = U^d + bU^b, \]

plus the utility of the supervisors (if any).

**Timing:** The timing of the decisions by market participants are as follows:

- Date 1-A: Congress sets the minimum capital requirement \( \underline{k} \).
• Date 1-B: A debt contract \((k, i)\) is offered by bankers.

• Date 1-C: Each banker borrows and chooses to invest in either a safe or a risky project.

• Date 2: Projects realize their return. Solvent bankers pay back \((1 - k)(1 + i)\) and consume the remaining profits. Insolvent bankers are liquidated.

### 1.5 Laissez-Faire Debt Contract

Let us first consider the case without any government intervention. That is, Congress neither implements any capital requirement, nor does it intervene in the market in any way.

It is noteworthy that if bankers offer to borrow less than \(\xi\), then depositors will be willing to lend at 0 interest rate. That is because regardless of bankers’ investments, depositors can always recover their deposits. However, I assume equity is expensive to bankers (i.e. \(\delta\) is sufficiently large) so that they would like to be able to borrow beyond \(\xi\). For this reason, if bankers’ projects fail, all of their assets will be seized by depositors.

In this section, I focus on debt contracts that guarantee investment in safe projects because they are socially desirable by (1.1) and (1.2). Although neither depositors nor bankers take social welfare into account in a laissez-faire environment, the result obtained in this section will provide a convenient benchmark for later analysis that explores how a benevolent Congress can improve welfare.

Bankers proposes a debt contract of the form \((k, i)\) that solves:

\[
\max_{k,i} U^b(k, i) = 1 - k + \frac{1}{\delta} \bar{r}\left[\xi - (1 - k)(1 + i)\right],
\]  
(1.4)
subject to:

\[ \frac{\Delta P}{\Delta r} \geq (1 - k)(1 + i), \quad (1.5) \]

\[ \bar{r}(1 - k)(1 + i) + (1 - \bar{r})\lambda \varepsilon \geq 1 - k. \quad (1.6) \]

\[ U^b(k, i) \geq 1 \quad \text{and} \quad \varepsilon \geq 1 + i. \quad (1.7) \]

The profit-maximization problem given by (1.4) indicates that bankers commit to pursue safe projects. The commitment is credible as long as the (simplified) Incentive Compatibility constraint given by (1.5) is met. Inequality (1.6) is depositors’ Participation Constraint, which ensures that depositors can at least expect to break even on their loan. Lastly, the first inequality in (1.7) is bankers’ Participation Constraint; whereas the second ensures that, given a bank’s project succeeds, the per-unit cost of borrowing that a bank has to pay does not exceed the per-unit return the bank gets.

Using Kuhn-Tucker conditions to solve for (1.4), subject to (1.5), (1.6) and (1.7), we have\(^\dagger\):

\[ 1 - k^{LF} = \bar{r} \frac{\Delta P}{\Delta r} + (1 - \bar{r})\lambda \varepsilon, \quad (1.8) \]

\[ i^{LF} = \frac{(1 - \bar{r})[\frac{\Delta P}{\Delta r} - \lambda \varepsilon]}{1 - k^{LF}}. \quad (1.9) \]

Bankers’ utility is:

\[ U^b(k^{LF}, i^{LF}) = \bar{r} \frac{\Delta P}{\Delta r} + (1 - \bar{r})\lambda \varepsilon + \frac{\bar{r}}{\delta}(\varepsilon - \frac{\Delta P}{\Delta r}) \]

A simple algebraic inspection will reveal that \( U^b(k^{LF}, i^{LF}) > 1 \) due to inequality (1.1). It follows that bankers are better off borrowing and investing in

\(^{\dagger}\)It can be verified that \( 1 - k^{LF} > \varepsilon \). That is, bankers are in fact borrowing beyond \( \varepsilon \) and risking liquidation.
safe projects than merely consuming their endowment in period 1.

Furthermore, bankers will not choose to borrow and invest in risky projects because the expected payoff from risky projects is less than 1, by \( (1.2) \).

We have the following proposition:

**Proposition 1.5.1.** *In a laissez-faire environment, the debt contract \((k^{LF}, i^{LF})\) given by \((1.8)\) and \((1.9)\) is a unique equilibrium. It also guarantees that all banks invest in safe projects.*

*Proof.* See discussion above. \(\square\)

Social welfare in this case is:

\[
W^{LF} = w + b \left[ \bar{r} \frac{\Delta P}{\Delta r} + (1 - \bar{r}) \lambda \varepsilon + \frac{\bar{r}}{\delta} (\varepsilon - \frac{\Delta P}{\Delta r}) \right]
\]  
(1.10)

### 1.6 Regulation and Supervision of Banks

#### 1.6.1 Supervision Without An Ex-post Bailout

Let us consider the case that Congress commits not to intervene in the market *ex post*. In particular, there will be no bailout when banks fail.

From the previous section, we see that a competitive equilibrium with only safe projects can be achieved; and that it yields a corresponding level of the expected social welfare \(W^{LF}\), given by \( (1.10) \). In this section, we explore whether Congress can bring social welfare to a higher level by putting in place a supervisory structure.

It is noteworthy that prudential regulation in the form of a minimum capital requirement is unwarranted in this case. That is because there is no transfer
from Congress, and hence, no excessive risk taking on bankers’ part. Any minimum capital requirement (different than $k^{LF}$) imposed by Congress will either be non-binding or too restrictive, which either has no effect on welfare or indeed reduces it. Thus, this section will consider supervision without capital requirement.

The following sections will then consider alternative arrangements for a supervisory structure. Accordingly, this section will show that supervisory competition as a result of supervisory shopping helps improve social welfare.

1.6.1.1 Incentive-based Supervision

Let us consider a supervisory structure with one supervisor. The supervisor possesses an auditing technology that can help identify risky banks. The effectiveness of the technology is dependent on the level of effort that the supervisor chooses to exert. The higher the effort, the better the supervisor can detect bankers’ choice of projects.

Let $e \in [0, 1]$ denote the supervisor’s effort level. For simplicity, let $e$ also be the probability that the supervisor finds out the choice of projects that any given bank has taken on. To generate $e$, the supervisory cost is given by $\psi(e)$ per bank, where $\psi(e)$ is strictly increasing and convex in $e$. Furthermore, $\psi(0) = 0$, $\lim_{e \to 0} \psi'(e) = 0$ and $\lim_{e \to 1} \psi'(e) = +\infty$.

The supervisor, whose reservation utility is 0, is required to examine all banks.

Alternatively, we can interpret $e$ as the fraction of banks that the supervisor (randomly) chooses to examine. In this interpretation, the supervisor finds out bankers’ choice of project with certainty. While both interpretations lead to the
same analytical results, the first interpretation will be assumed.

In addition, the supervisor is given an orderly liquidation authority. If a bank is found pursuing a risky project, the supervisor places the bank under receivership and takes on the duties of selling the bank’s asset. I assume that the supervisor can redeem the assets for the original amount of investment. The recovered capital becomes a part of the public fund, which later pays out to depositors. (As a punishment for taking risky projects, bankers do not get their equity back.) The described supervisory mechanism and asset recovery, while greatly simplified, are similar to orderly liquidation process in real world bank resolutions, including the Orderly Liquidation Authority of the FDIC in the US and the Single Resolution Mechanism proposed by the European Commission for the forthcoming Banking Union in the Eurozone.

I assume that the supervisor’s findings about banks is private. Thus, the supervisor can declare safe a bank that she finds risky or that she does not find any information. In this way, the supervisor can choose not to close a risky bank. However, since a bank can bring charges against the supervisor if it is wrongfully closed (which is assumed to be costly to the supervisor), the supervisor shuts down a bank only if she has evidence that the bank is in fact pursuing a risky project.

Also to simplify the analysis, I assume that the supervisor sets $e$ before bankers make investment decisions. The assumption can be justified by recognizing that the supervisory effort $e$ reflects slow-moving variables such as the supervisor’s human capital or capacity. In both cases, $e$ is established early and difficult to change in comparison to bankers’ choice of projects. Without this assumption, dynamic inconsistency may arise as the supervisor may re-optimize
her effort ex post.

From the perspective of Congress, the supervisory effort $e$ is not verifiable, and hence, cannot be used as a contractual condition. Congress then designs a transfer payment $s$ conditional on two verifiable pieces of information: (i) the number of supervised banks $\hat{b}$, for $0 \leq \hat{b} \leq b$, and (ii) the number of bank closures $b^c$, for $0 \leq b^c \leq \hat{b}$. Since there is only one supervisor at this point, we have $\hat{b} \equiv b$.

I assume that the transfer is financed by a lump-sum tax on depositors.

With the presence of a supervisory structure, the timing of the game is as follows:

- Date 1-A: Congress offers a supervisory contract $s(b^c)$.
- Date 1-B: If the supervisor accepts the contract, she chooses to exert effort $e$.
- Date 1-C: A debt contract of the form $(k, i)$ is offered by bankers. Each banker borrows and chooses to invest either in safe or risky projects.
- Date 1-D: The supervisor examines all banks. Risky banks that are found are placed into receivership and dissolved.
- Date 2: Projects realize their return. Solvent banks pay back $1 - k$ and consume the remaining profit. Insolvent banks are resolved by Congress. Taxes are levied.
1.6.1.2 Verifiable Supervisory Effort

In this section, let us assume that the supervisor’s effort is verifiable ex post by Congress. A vital contrast will then become apparent in the next section, as we resume the assumption that effort is not verifiable.

In this case, Congress can induce a supervisory effort level $e$ by simply paying the supervisor $s = b\psi(e)$.

Given a supervisory effort level $e$, bankers offer a debt contract $(k, i)$ that solves the same profit maximization problem given by (1.4), subject to Depositors’ Participation Constraint (1.6) and inequality (1.7) in Section 1.5. Bankers’ Incentive Compatibility Constraint in this case is:

$$
\bar{r}[\varepsilon - (1 - k)(1 + i)] \geq (1 - e)\bar{r}[R - (1 - k)(1 + i)] \\
\Leftrightarrow \frac{\bar{r}[\varepsilon - (1 - e)\bar{r}R]}{IC(e)} \geq (1 - k)(1 + i). \tag{1.11}
$$

Note that for $e = 0$, inequality (1.11) is reduced to the Incentive Compatibility condition (1.5) in Section 1.5. Furthermore, it is easy to verify that the left hand side of (1.11), $IC(e)$, is increasing and concave in $e$.

Using Kuhn-Tucker conditions to solve for (1.4), subject to (1.6), (1.7) and (1.11), for a given level of $e$, we have:

$$
1 - k(e) = \bar{r}IC(e) + (1 - \bar{r})\lambda \varepsilon, \tag{1.12}
$$

$$
i(e) = \frac{(1 - \bar{r})[IC'(e) - \lambda \varepsilon]}{1 - k(e)} \tag{1.13}
$$

Congress chooses $e$ to maximize social welfare, given by:

$$
\max_e W(e) = w - b\psi(e) + b\left[\bar{r}IC(e) + (1 - \bar{r})\lambda \varepsilon + \frac{\bar{r}}{\delta}(\varepsilon - IC(e))\right]
$$
The first order condition is:

\[ W'(e) = 0 \]

\[ \Rightarrow \psi'(e) = (1 - \frac{1}{\delta})rIC'(e). \quad (1.14) \]

Since \( W''(e) \) is strictly negative, there exists at most one \( e^* \) such that equation (1.14) is satisfied. Furthermore, an interior solution \( 1 > e^* > 0 \) exists because \( \lim_{e \to 1} \psi'(e) = +\infty \).

We have the following proposition:

**Proposition 1.6.1.** If supervisory effort is verifiable ex post, Congress can implement a supervisory effort \( e^* \) and achieve a level of social welfare that is strictly higher than \( W^{LF} \), given by (1.10) in Section 1.5. Furthermore, \( k(e^*) < k^{LF} \) and \( i(e^*) > i^{LF} \).

**Proof.** First, as discussed above, Congress can implement \( e^* \) by giving the supervisor a contract that pays \( b\psi(e^*) \) if and only if \( e = e^* \).

Second, it is straightforward to see that Congress can always achieve \( W^{LF} \) by implementing \( e = 0 \). That is because \( W(0) = W^{LF} \).

Third, let us note that \( W'(0) \) is strictly positive because \( \psi'(0) = 0 \) and \( IC'(0) > 0 \). It follows that \( e = 0 \) cannot be a solution to Congress’ welfare maximizing problem. Hence, for \( e^* \) that satisfies the first order condition (1.14), it must be the case that \( e^* > 0 \) and \( W(e^*) > W(0) = W^{LF} \).

Finally, it is straightforward to verify that \( k'(e) \) is strictly negative while \( i'(e) \) is strictly positive, which proves the last claim.

\( \square \)
In comparison to the previous section, we see that supervision can act as a beneficial substitute for capital. The more efficient the supervisor, the less capital bankers are required to bring. Balancing between the cost of supervision and cost of equity (via (1.14)), Congress can therefore help bankers raise more deposits and bring social welfare to a higher level. Despite the fact that bankers have to borrow at a higher interest rate, they enjoy a higher overall expected profit by being able to save on equity.

Nonetheless, the result is obtained under the assumption that supervisory effort is verifiable. In the next section, we will see that it is difficult to motivate a single self-interested supervisor to exert positive non-verifiable effort. Thus, more than one supervisor will be required.

1.6.1.3 Unverifiable Supervisory Effort

Now, let us suppose that the supervisor’s effort is not verifiable by Congress. It follows that the supervisory contract cannot be directly contingent on her effort.

Recall that the supervisory contract will take the form of a transfer \( s(\hat{b}, b^c) \), contingent on the number of supervised banks \( \hat{b} \) by the supervisor and the number of banks found to be risky, \( b^c \). Since there is only a single supervisor, \( \hat{b} = b \).

We then have the following proposition:

**Proposition 1.6.2.** If supervisory effort is not verifiable and the supervisor’s contract can be made contingent only on the number of closed banks, there is no contract that can induce positive effort from a single self-interested supervisor and lead to a safe banking environment.
Proof. In the previous section, we have seen that a debt contract of the form 
\((k(e), i(e))\) given by (1.12) and (1.13) is necessary and sufficient to induce a safe 
banking environment. That is, for a given level of \(e\), bankers and depositors 
will enter a debt contract that guarantees investment in safe projects.

Now suppose that Congress has given the self-interested supervisor an incen-
tive contract \(s(b, b^c)\) at Date 1-A. In expectation of a safe banking environment, 
it must be the case that \(b^c\) – the number of risky banks the supervisor is required 
to find – is equal to 0. Hence, the contract offered by Congress in this case is a 
fixed transfer payment \(s(b, 0) \geq 0\).

Recall that the supervisor’s effort is not verifiable. That means the super-
visor can act as if she has exerted sufficiently high effort level and catalyze safe 
banking. In other words, the supervisor fulfills the contract by declaring that 
she has found no risky banks regardless of her actual effort.

The supervisor therefore chooses \(e\) to maximize:

\[
\max_{e} U^s(e) = s(b, 0) - b\psi(e).
\]

Consequently, the supervisor always chooses 0 effort.

The result in Proposition 1.6.2 is intuitive. That is, to be efficient and 
able to deter risky banks, a supervisor has to exert high effort and bear the 
respective supervisory cost. However, an efficient supervisor does not close any 
banks, because all banks are safe. It follows that a self-interested supervisor, 
while minimizing her cost, will mimic the action of an efficient one and closes no 
bank. Hence, whenever Congress gives the supervisor a contract in expectation 
of safe banking, the self-interested supervisor will always shirk. Congress is
unable to motivate the single supervisor to work; and social welfare remains as if there is no supervision.

1.6.1.4 Multiple Supervisors and A Race to the Top

Section 1.6.1.2 and 1.6.1.3 reveals important features of supervision. On the one hand, supervision, when effort is verifiable, can help improve social welfare by allowing the capital requirement to be lower. Bankers can borrow more, and save on the cost of equity. On the other hand, supervision becomes very difficult to incentivize when the effort of the supervisor is not verifiable.

In this section, I propose a setting with two identical supervisors to solve the above dilemma. We will see that the supervisors are motivated to compete with one another, and hence, to exert high effort level.

With multiple supervisors, let bankers have the ability to select a supervisor before investing. To simplify the analysis, I assume that if bankers are indifferent between the two supervisors, each banker randomizes his choice. Hence, each supervisor will expect to get \( b/2 \) banks.

Let \( e_1 \) and \( e_2 \) denote the effort levels of supervisor 1 and 2, respectively.

Similar to the analysis in Section 1.6.1.3, bankers who have chosen supervisor \( j \) with effort \( e_j \) solve the profit maximization problem given by (1.4), subject to (1.6), (1.7) and (1.11). The debt contract under supervisor \( j \) will be:

\[
1 - k(e_j) = \bar{r}IC(e_j) + (1 - \bar{r})\lambda \xi,
\]

\[
i(e_j) = \frac{(1 - \bar{r})[IC(e_j) - \lambda \xi]}{1 - k(e_j)}
\]

Bankers’ expected utility is:

\[
U^b(k(e_j), i(e_j)) = \bar{r}IC(e_j) + (1 - \bar{r})\lambda \xi + \frac{\bar{r}}{\delta}(\varepsilon - IC(e_j))
\]
Since \( IC(e_j) \) is strictly increasing, it is easy to verify that \( U^b(k(e_j), i(e_j)) \) is also increasing in \( e_j \). The intuition for this result is that the higher the supervisory effort, the easier it is for bankers to commit to safe investment. Hence, bankers will be able to increase their leverage by raising more deposits and using less of their own capital. That, in turn, raises bankers’ overall expected utility.

As a consequence, bankers’ preference over supervisors is simple: they prefer a supervisor with high supervisory effort.

We have the following proposition:

**Proposition 1.6.3.** In the presence of two supervisors, Congress can induce both supervisors to exert any effort level \( e_C \).

**Proof.** Let \( b_j \) denote the number of banks that supervisor \( j \) supervises.

Consider a contract that pays:

\[
s(b_j, 0) = b_j \psi(e_C),
\]

for \( j \in \{1, 2\} \).

First, neither supervisors will exert effort beyond \( e_C \). That is because for any positive number of banks a supervisor gets, the contract compensates the supervisor only \( \psi(e_C) \) per bank. A self-interested supervisor will, therefore, choose her effort level to be less than or equal \( e_C \).

Second, if one supervisor exerts effort strictly below \( e_C \), it is optimal for the other to choose her effort to be marginally higher. Since bankers strictly prefer a supervisor with a higher effort, the high-effort supervisor will get all the banks. The low-effort supervisor will be left with no bank, and hence, no income. Consequently, it is never optimal for the two supervisors to exert effort strictly below \( e_C \).
Finally, not only the outcome in which both supervisors exert $e_C$ is a Nash equilibrium, it is the unique one. If one supervisor raises her effort, she gets all the banks but her supervisory cost exceeds the contract payment. If she lowers her effort, she loses all the banks, and therefore does not receive any income. Furthermore, the arguments above show that there cannot be any other equilibrium.

Proposition 1.6.3 has shown that by having multiple supervisors, Congress can induce healthy competition among them. This is a consequence of supervisory shopping. As bankers “vote with their feet” to elect the supervisor with the highest level of effort, Congress rewards supervisors by the number of banks they receive. This indirectly makes the effort verifiable. Congress can then motivate any level of supervisory effort it desires. The welfare-maximization problem that Congress faces now becomes the same as in Section 1.6.1.2, where supervisory effort is verifiable.

The optimal contract in this case is similar to the funding schemes for many state and federal banking supervisory agencies in the US, including the oldest federal banking authority—the OCC. Before the creation of the OCC in 1863, the banking industry was unstable across the nation since half of all banks failed within a few years of inception (Komai and Richardson, 2011). For many decades thereafter, the OCC and state banking authorities competed vigorously to attract new banks by continually adjusting their regulatory and supervisory standards (White, 1982). This may explain why the American banking industry both expanded rapidly and became much more stable during this period (White, 2009).
Taking a broader perspective, the result provided in this section can also be applied to competition among stock exchanges. To attract firms, exchanges often compete to raise their requirements, thereby make listed firms safer and more valuable to investors. Brunnermeier, Huddart, and Hughes (1999) argue the same point. Meanwhile, Doidge, Karolyi, and Stulz (2004) shows that high disclosure standards in the US positively affect foreign firms’ decision to list in the US. In other words, firms have strong incentives to list in places where tough standards and requirements can increase firms’ value. Exchanges, therefore, enter a “race to the top” to attract firms.

1.6.2 Capital Requirement and Bailout Guarantee

1.6.2.1 Regulation Without Supervision

In previous sections, we have analyzed different supervisory environments under the assumption that there is no government intervention. In particular, Congress can commit not to bail out the banks when their projects fail.

The commitment, however, is not credible ex post. That is because even in a safe banking environment, a low-return shock can make \( b(1 - \bar{r}) \) banks insolvent. Since depositors can recover only \( \lambda \xi \) when banks declare bankruptcy, the deadweight loss to social welfare is \( b(1 - \bar{r})(1 - \lambda)\xi \). Thus, an ex-post welfare-maximizing Congress will be inclined to save the insolvent institutions from bankruptcy.

Let a bailout be specified by a transfer to failing banks, financed by a tax on all depositors. Since an ex-post bailout is always optimal, market participants will make decision as if the bailout is guaranteed.

Besides, Congress will also act as a conservator for failed banks. Particularly
at Date 2, Congress places insolvent banks under receivership, takes control of all of their remaining assets and assumes all of their debt. Then, Congress imposes a tax on all depositors so that, together with insolvent banks’ assets, it can repay in full the debt of those insolvent banks.

The process of bank resolution described above is common in practice, particularly similar to the proposal by the European Commission for the future banking union in Europe and to the liquidation process the FDIC currently uses in the United States. Although in both cases, membership fees are imposed on banks and used to resolve bank failures, bank resolutions are ultimately backed by taxpayers.

With deposits in effect being guaranteed by Congress, depositors are willing to lend (unlimitedly) at 0 interest rate. If there is no minimum capital requirement, bankers will contribute 0 equity capital. Hence, they enjoy the profit when projects succeed, but lose nothing if their projects fail. By (1.3), every atomistic banker will choose to pursue a risky project. The eventual cost of the bailout will exceed the total profit earned by all solvent banks because of (1.2). As a result, social welfare will fall below the total of the initial endowments.

To ensure bankers pursue only safe projects (since safe projects is socially desirable by (1.1)), Congress will need to implement a capital requirement that satisfies:

\[ \bar{r}(\varepsilon - (1 - k)) \geq r(R - (1 - k)) \]

\[ \Rightarrow k \geq 1 - \frac{\Delta P}{\Delta r}. \]

(1.15)

For a given \( k \), bankers choose \( k \geq \bar{k} \) to maximize their profit. It is straightforward to show that bankers will always pick \( k = \bar{k} \).
The expected tax needed for a bailout is:

\[ \tau = b(1 - \bar{r})[(1 - k) - \varepsilon]. \]

Congress chooses \( k \) to solve:

\[ \max_k W(k) = w - \tau + b \left[ 1 - k + \frac{\bar{r}}{\delta}(\varepsilon - (1 - k)) \right], \]

subject to (1.15).

Thus, the capital requirement:

\[ k^0 = 1 - \frac{\Delta P}{\Delta r} \] (1.16)

maximizes social welfare.

Social welfare in this case is:

\[ W^k = w + b \left[ (1 - \frac{1}{\delta})\frac{\Delta P}{\Delta r} + (1 - \bar{r})\varepsilon \right] + \frac{1}{\delta} [\bar{r}\varepsilon + (1 - \bar{r})\varepsilon] \] (1.17)

Note that:

\[ W^k - W^{LF} = b(1 - \bar{r})(1 - \lambda)\varepsilon, \]

which is the exact amount of the deadweight loss when there is no bailout.

We have the following proposition:

**Proposition 1.6.4.** When deposit is guaranteed, the capital requirement \( k^0 \), given by (1.16), is necessary and sufficient to induce safe banking and maximize social welfare. Social welfare \( W^k \), given by (1.17), is higher than welfare \( W^{LF} \) in a laissez-faire environment, given by (1.10).

**Proof.** See discussion above.
With deposit guarantees, social welfare is improved. Bankers increase their expected profit because, in comparison to the previous case, they can borrow more \((k^0 < k^L_F)\) and do so at 0 interest rate.

For the remainder of the analysis, a bailout is assumed to be guaranteed. I will now consider the impact of supervision and supervisory shopping in the new environment.

### 1.6.2.2 A Single Supervisor

First, let us consider the case where supervisory effort is verifiable as in Section 1.6.1.2. The difference in this section in comparison to Section 1.6.1.2 is that, as a result of the guaranteed bailout, Congress will maximize welfare by simultaneously choosing the minimum capital requirement, the tax on solvent banks and the supervisory effort. Furthermore, the interest rate remains at 0.

Recall that:

\[
IC(e) = \frac{\bar{r} \varepsilon - [1 - e] \bar{r} R}{\bar{r} - [1 - e] \bar{r}}
\]

Then, Congress solves the following maximization problem:

\[
\max_{k,e} W^\text{bailout}(k,e) = w - \tau + b \left(1 - k + \frac{\bar{r}}{\delta} \varepsilon - (1 - k)\right),
\]

subject to:

\[
IC(e) \geq 1 - k,
\]

\[
\tau = b[\psi(e) + (1 - \bar{r})(1 - k - \varepsilon)].
\]

Inequality (1.19) is bankers’ Incentive Compatibility Constraint, ensuring that bankers will pursue safe projects. The tax that is needed to cover the supervisory and bailout costs is given in (1.20).
Kuhn-Tucker Conditions yield:

\[ \psi'(e) = \bar{r}(1 - \frac{1}{\delta})IC'(e) \quad (1.21) \]

\[ IC(e) = 1 - k. \quad (1.22) \]

Similar to Section 1.6.1.2, equation (1.21) has a unique solution at \( e = e^* > 0 \). It follows that the capital requirement in this case is \( k^* = 1 - IC(e^*) \).

Note that the optimal supervisory effort in this case is the same as in the absence of bailouts. In other words, the availability of government guaranteed bailouts does not directly alter bankers’ appetite for risk. To enforce safe banking, the supervisor is merely required to exert the same effort level as before. Furthermore, the bailouts help bankers save on equity \( k^* < k^0 \), see Proposition 1.6.5. That is to say when supervisory effort is verifiable, bailouts do not lead to the usual moral hazard problem, but indeed helps improve social welfare by eliminating the need for bankruptcy and the associated cost.

We have the following proposition:

**Proposition 1.6.5.** Given that a bailout is guaranteed and that supervisory effort is verifiable, Congress can implement a combination of capital requirement and supervisory effort \((k^*, e^*)\) that achieves a strictly higher level of social welfare than \( W^k \), given by (1.17) in Section 1.6.2.1. Furthermore, \( k^* < k^0 \).

**Proof.** First, recall that \( IC(e) \) is strictly increasing in \( e \). Since \( IC(0) = \frac{\Delta P}{\Delta r} \) and \( e^* > 0 \), it follows that \( IC(e^*) > \frac{\Delta P}{\Delta r} \). Hence, \( k^* < k^0 \).

Second, note that \( W^{\text{bailout}}(k^0, 0) = W^k \). Since \( W^{\text{bailout}}(k, e) \) is uniquely maximized at \((k^*, e^*)\), where \( e^* > 0 \) and \( k^* < k^0 \), it follows that \( W^{\text{bailout}}(k^*, e^*) > W^k \). \( \square \)
Thus, similar to Section 1.6.1.2, supervision, when supervisory effort is verifiable, can help bring social welfare to a higher level. Capital requirement is lower; and bankers enjoy higher utilities. However, Proposition 1.6.2 in Section 1.6.1.3 continues to apply. That is, if supervisory effort is not verifiable, the supervisor will always maximize her expected utility by setting effort to 0. It follows that a supervisory structure with a single self-interested supervisor does not improve social welfare.

1.6.2.3 Two Supervisors and A Race to the Bottom

Recall that Section 1.6.1.3 and 1.6.1.4 have shown that a supervisory system requires more than one self-interested supervisor to be effective. That is primarily due to the fact the supervisory effort is not verifiable. A single supervisor, then, has a strong incentive to minimize her effort. With multiple supervisors, Congress can motivate the supervisors to compete and thereby achieve a high level of social welfare. One important reason behind this result is that, in a bailout-free environment, bankers strongly prefer supervisors to exert high effort so that they can reduce their equity contribution.

Unfortunately, we will see in this section that the preference of bankers changes when a bailout is guaranteed. Bankers are able to borrow at 0 interest rate while a minimum capital requirement is always binding. Bankers will, therefore, seek out a low effort supervisor so that they may take on risky projects and increase their expected profit.

To see this, let us consider the expected profit on safe and risky project,
given a capital requirement $k$ and supervisory effort $e$:

$$\Pi_s = \bar{r}[\varepsilon - (1 - k)],$$

$$\Pi_r(e) = (1 - e)R[1 - (1 - k)] = [1 - e]\Pi_r.$$

First, let us focus on $k < \underline{k}^0$, for $\underline{k}^0$ given by (1.16) in Section 1.6.2.1. That is because otherwise, $\Pi_s \geq \Pi_r(e), \forall e \geq 0$, and supervision becomes irrelevant.

Second, for $k < \underline{k}^0$, let $\bar{e}(k)$ be such that:

$$\Pi_s = (1 - \bar{e}(k))\Pi_r. \quad (1.23)$$

It follows that for $e > \bar{e}(k)$, bankers will prefer safe projects. Since the expected profit from safe projects, $\Pi_s$, does not depend on $e$, bankers are indifferent between supervisors if they both exert effort higher than $\bar{e}(k)$.

Meanwhile, for $e < \bar{e}(k)$, bankers will prefer risky projects. Furthermore, the lower the effort, the higher the expected profit from risky projects, $\Pi_r(e)$, becomes. Given a choice between two supervisors whose effort levels are distinct (but both remains below $\bar{e}(k)$), bankers will prefer the supervisor with lower effort.

Now, given a contract $s(\hat{b})$ and a capital requirement $k$, let us consider supervisors’ decision.

For the contract to be more realistic, I assume that the contract payment must satisfy the following condition:

$$s(\hat{b}, 0) \geq \psi(\bar{e}(k))\hat{b}, \forall \hat{b} \in [0, b], \quad (1.24)$$

where $\bar{e}(k)$ is given by (1.23).
The intuition for the assumption is as follow. Given a capital requirement $k$, $\bar{e}(k)$ is the minimum level of effort that induces bankers to take safe projects. Thus, the supervisory contract must at least cover the cost of supervision for any bank allocation across supervisors based on $\bar{e}(k)$. This is to prevent the case in which a supervisor exerts the effort level necessary for safe banking but is selected by more banks than her contract would cover. Although this case would not exist within the equilibrium, the out-of-equilibrium assumption (1.24) is needed for the following proposition.

**Proposition 1.6.6. A Race to the Bottom:** Given Assumption (1.24), there is no supervisory contract that leads to positive supervisory effort in an environment with a guaranteed bailout.

**Proof.** Let $e_j$ and $b_j$ respectively denote the effort level and the number of banks supervisor $j$ has, for $j \in \{1, 2\}$, given that a contract $s(\hat{b}, b^c)$ has been offered and accepted. In expectation of safe banking, supervisors should not be required to close any risky banks. Hence, $b^c = 0$.

The expected utility of supervisor $j$, for $j \in \{1, 2\}$ is:

$$U_j(e_j) = s(b_j, 0) - b_j \psi(e_j).$$

Recall that for a capital requirement $k$, bankers are indifferent between the two supervisors for $e \geq \bar{e}(k)$ and prefer low effort supervisors otherwise. It follows that given the other supervisor’ effort, a supervisor can always reduce her own effort and expect to receive at least as many banks as before. By Assumption (1.24), the supervisory payment must be non-decreasing in the number of supervised banks.
Hence, each supervisor maximizes her expected utilities by choosing 0 effort.

In Appendix A.1, I discuss the case where Assumption (1.24) does not hold. There is an optimal contract that can reveal supervisors’ effort, and as a result, help achieve the first-best welfare level. Similar to the idea proposed by Shleifer (1985), Congress can offer a contract that rewards one supervisor based on the number of banks that the other supervisor has (“yardstick competition”). Given bankers’ incentive in the presence of a guaranteed bailout, supervisors under this type of contract will compete in raising their effort, not to be more attractive to banks, but to “shift” banks away from them. However, it is unlikely that such a contract can be implemented. That is because while supervisors may strive to appear as more or less attractive to banks, banks have the final say on who their supervisor is. Thus, at the appropriate level of effort required for safe banking, the supervisory contract should always give to the supervisors enough resources to supervise banks for any allocation of banks across supervisors. (This is also the exact reason for Assumption (1.24).)

Given Assumption (1.24), Proposition 1.6.6 holds true and echoes the same result as Proposition 1.6.2 in Section 1.6.1.3. That is, when supervisors are self-interested and their effort is not verifiable, it is difficult to motivate them to work. It is particularly not possible to induce positive effort in this section because bankers’ preferences for supervisors have change. In Section 1.6.1.4 without a guaranteed bailout, Congress relies on bankers’ inclination toward high effort supervisors to induce the supervisors to work. In this section, however, a guaranteed bailout make it easier for bankers to raise deposits, which in
turn make risky projects more attractive. As a consequence, bankers now seek low effort supervisors.

In such an environment, it is optimal for Congress to rely entirely on strict regulation, namely a high level of capital requirement, to achieve safe banking and maximize social welfare. Specifically, Congress implements the minimum capital requirement $k^0$, given by (1.16) in Section 1.6.2.1.

1.7 Extensions

1.7.1 CoCos and The Best of Both Worlds

Section 1.6.1 and 1.6.2 have provided two contrasting view of supervisory shopping in the banking sector. While the former shows that shopping leads to effective supervision, the result depends on a no-bailout commitment by Congress. This commitment is not credible, however, because the deadweight loss of bankruptcy lowers social welfare. Correspondingly, the latter section shows that supervisory shopping makes it difficult to incentivize supervision in the presence of bailouts.

In this section, I examine the case in which bankers can issue CoCos and argue that the use of CoCos can help achieve the best of both worlds: efficient supervision and the avoidance of the deadweight loss associated with bankruptcy. I focus the attention on CoCos because they are debt-equity hybrid capital securities that can absorb losses when issuing banks are in distress. However, non-CoCos debt, including subordinated debt, can typically absorb losses only after issuing banks have fallen into liquidation or bankruptcy. Albul, Jaffee, and Tchistyi (2013), Pennacchi (2010) and Glasserman and Nouri (2012) offer
formal treatments of CoCos, whereas Pazarbasioglu, Zhou, Le Leslé, and Moore (2011) and Avdjiev, Kartasheva, and Bogdanova (2013) discuss the economic rationale and recent changes in the CoCos market.

Following the criterion outlined by Avdjiev, Kartasheva, and Bogdanova (2013), I assume CoCos have the following structure:

1. Each banker issues a bond in exchange for $B_{CoCos}$ from depositors at Date 1
2. The bond promises to pay $B_{CoCos}(1 + i_{CoCos})$ to depositors at Date 2
3. At Date 2, if bankers fail to repay the depositors, the bond is converted into 100% equity.

While the first two conditions are typical of any bond, the last one that specifies a mechanical trigger and a loss absorption mechanism represents the unique feature of CoCos. Note that the trigger and the loss absorption mechanism can vary greatly in reality; for simplicity, I choose this particular form. The qualitative results presented in this section are not affected by this choice.

Bankers’ liabilities now consist of equity capital, CoCos and deposits. Let $k$ continue to denote equity capital; $K$ denote the total of equity and CoCos $(K = k + B_{CoCos})$. It follows that the amount of deposits is $1 - K$.

Regulatory capital now is characterized by $K$, which is the minimum total capital requirement, comprising of both equity capital and CoCos (Tier 2 capital).

Before proceeding, let us recall that in Section 1.6.1, the highest level of social welfare that is attainable is $W(e^*)$. While the optimal supervisory effort
is $e^*$, bankers contribute $k(e^*)$ of their own equity capital. The result, however, depends on a no-bailout commitment that is not credible due to a deadweight loss of bank liquidation.

Similarly, in Section 1.6.2, the highest level of social welfare, $W_{bailout}(k^*, e^*)$, is attainable under the assumptions that a bailout is guaranteed and supervisory effort is verifiable. However, I then show that when effort is not verifiable, $(k^*, e^*)$ is not feasible.

We then have the following proposition:

**Proposition 1.7.1.** If Congress imposes a minimum total capital requirement $K^* = 1 - \varepsilon$, then:

- The no-bailout commitment is credible;
- The optimal supervisory effort $e^*$ is attainable by having two supervisors;
- Bankers borrow with $k = k(K^*, e^*)$ equity capital such that $k(K^*, e^*) < k^*$
- Finally, social welfare is higher than $W(e^*)$ and $W_{bailout}(k^*, e^*)$.

**Proof.** See Appendix A.2

The intuition of Proposition 1.7.1 is as follow. When Congress implements $K^*$, bankers can raise only $\varepsilon$ in deposits. The rest is in equity capital and CoCos. If there is a crisis and their projects fail, CoCos convert into equity. Bankers can still cover their deposit liabilities and therefore avoid declaring bankruptcy. It follows that the no-bailout commitment is credible. Furthermore, without an expected bailout, bankers seek strong supervisors to raise their leverage. Similar to Section 1.6.1.4, supervisory shopping then leads to a race to the top among
supervisors. Congress, therefore, finds it optimal to have two supervisors and to implement supervisory effort \( e^* \).

Because there is no deadweight loss from liquidation in this case, social welfare is higher than \( W(e^*) \) in Section 1.6.1. Meanwhile, because effective supervision can be implemented and there is no social cost of a bailout, social welfare is higher than \( W_{\text{bailout}}(k^*, e^*) \) in Section 1.6.2. Via the use of CoCos, we therefore have the best of both worlds!

### 1.7.2 Shopping in An International Context

To study shopping in an international context, I extend the model to two countries. As each country has its own regulatory regime, bankers are free to choose their chartering country. This extension reflects the situation in Europe as argued in Section 1.2. To simplify the analysis, I assume that only one of the two countries has a banking sector, and that the other country receives a share of bankers’ profit when these banks register there. These assumptions lead to the same qualitative results as when bankers live in both countries and are free to move.

Formally, let us consider two countries, \( G \) and \( L \). Let each country have its own set of depositors and Congress as described in Section 1.4. For simplicity, I assume that depositors do not move. The assumption can be rationalized by letting depositors have a high cost of relocation.

Moreover, I assume that only country \( G \) has a set of bankers as described in Section 1.4. Bankers are to raise deposits and invest inclusively in country \( G \); however, they have the option of registering the business in either country \( G \) or country \( L \). By registering with one of the two countries, bankers are subject
to regulation and supervision of that country alone. If a banker registers her business in country \( L \), an \( \alpha \) portion of her utility is counted toward country \( L \)'s welfare, for \( 0 < \alpha < 1 \). The remaining portion is then counted toward country \( G \)'s welfare. If a banker is indifferent between the two countries, I assume that she will stay in her native country - country \( G \).

Akin to previous analysis, I consider the two-country setting in both bailout-free and bailout-guaranteed environments. In the bailout-guaranteed environment, I assume that both countries share the cost of bailing out bankers who register in country \( L \). This assumption reflects a political economy approach of this paper: while the 1994 EU Directive on Deposit-Guarantee Schemes dictates that the home country of an international bank (in this case, country \( L \)) has the responsibility to provide the insurance for the bank’s deposit, cross-country fiscal transfers are usually untimely or altogether avoided, causing the host country (in this case, country \( G \)) having to foot the bill in order to avoid contagion and systemic failures (Schoenmaker and Oosterloo, 2007; Goodhart and Schoenmaker, 2009). The burden-sharing rule in the case of bank resolution is as follows: for every bank that fails under country \( L \)'s jurisdiction, country \( L \) is responsible for a fraction \( \beta \) of the cost of bailing out that bank, for \( 0 < \beta < 1 \).

From the analysis in previous sections, the only relevant cases here include a bailout-free environment with supervision and a bailout-guaranteed environment without supervision. That is because in a bailout-free environment, the market is capable of reaching the safe banking equilibrium without capital requirement. The principal, at the same time, can implement any supervisory effort (Proposition 1.6.3). In a bailout-guaranteed environment, however, capital requirement is necessary. Unfortunately, supervision cannot be implemented
due to bankers’ preference towards weak supervision and self-interested supervisors’ non-veritable effort (Proposition 1.6.6).

We will now consider the two cases in turn.

1.7.2.1 Supervisory Competition Across Countries

In a bailout-free environment, we see in Section 1.5 that the market is capable of reaching the safe banking equilibrium without capital requirement. At the same time, Congress can employ multiple supervisors and motivate any level of supervisory effort as if effort was verifiable (Proposition 1.6.3). Thus, without loss of generality we proceed in this case with the assumption that effort if verifiable. By the Revelation Principle, Congress in each country can directly choose the level of supervisory effort that the domestic supervisors would exert.

Recall Section 1.6.1.2, we know that for a given supervisory effort $e$, bankers offer depositors $(k(e), i(e))$ in the bailout-free environment such that:

$$1 - k(e) = \bar{r}IC(e) + (1 - \bar{r})\lambda_\varepsilon,$$

$$i(e) = \left(1 - \bar{r}\right)\left[IC(e) - \lambda_\varepsilon\right] \frac{1 - k(e)}{1 - k(e)},$$

for $IC(e)$ given in (1.11).

Bankers’ utility is then:

$$U^b(k(e), i(e)) = \bar{r}IC(e) + (1 - \bar{r})\lambda_\varepsilon + \frac{\bar{r}}{\delta}(\varepsilon - IC(e)).$$

Now, consider $\hat{e}$ such that:

$$\bar{r}IC(\hat{e}) + (1 - \bar{r})\lambda_\varepsilon = 1.$$  \hspace{1cm} (1.25)

Note that $k(\hat{e}) = 0$. Intuitively, $\hat{e}$ is the level of supervisory effort that
enable bankers to borrow without contributing any capital. Because $\bar{r}IC(1) + (1 - \bar{r})\lambda\xi = \bar{r}\xi + (1 - \bar{r})\lambda\xi > 1$ by (1.1), it follows that $1 > \hat{e}$.

Let us consider bankers’ preference over chartering country. First, it is straightforward to verify that bankers’ utility is increasing with the supervisory effort, as long the effort does not exceed $\hat{e}$. That is because the better the supervision, the lower the capital bankers are expected to contribute to the projects. Accordingly, bankers prefer the country with better supervision. Second, if both country have supervisory effort exceeding $\hat{e}$, bankers do not have to contribute any capital, and hence, will be indifferent between the two countries. In that case, all bankers remain in country $G$.

Let the subscripts $G$ and $L$ denote the respective countries.

Given $e_G$, social welfare of country $L$ is given by:

$$W_L(e_L) = \begin{cases} w - b\psi(e_L) + \alpha bU^b(k(e_L), i(e_L)) & \text{if } e_L > e_G \text{ and } \hat{e} > e_G, \\ w & \text{otherwise.} \end{cases} (1.26)$$

Let $\bar{e}$ be such that:

$$\alpha U^b(k(\bar{e}), i(\bar{e})) = \psi(\bar{e}). (1.27)$$

Intuitively, $\bar{e}$ is the threshold beyond which supervisory cost in country $L$ exceeds the share of bankers’ utility that the country receives. From the perspective of country $L$, it is never beneficial to have supervisory effort exceeding $\bar{e}$. Also, $1 > \bar{e}$ because $\lim_{e \to 1} \psi'(e) = +\infty$.

Thus, country $L$ can maximize its welfare as follows:

- If $e_G \geq \min\{\hat{e}, \bar{e}\}$, no banker needs to register with country $L$. $W_L(e_L) \equiv w$ for all $e_L$.  

46
• If $e_G < \min\{\hat{e}, \bar{e}\}$, country $L$ attracts all banks and maximize welfare by setting $e_L$ marginally higher than $e_G$, satisfying $e_G < e_L \leq \min\{\hat{e}, \bar{e}\}$.

Country $L$'s policy as described above indicates that the country will always try to provide better supervision than country $G$ as long as the supervisory effort in country $G$ is below the threshold $\min\{\hat{e}, \bar{e}\}$.

To see how country $G$ respond, recall that the optimal supervisory effort for country $G$ in the absence of country $L$ is $e^*$ (see Section 1.6.1.2).

If $e^* \geq \min\{\hat{e}, \bar{e}\}$, country $G$ can simply set $e_G = e^*$. That guarantees all bankers will remain in country $G$ and maximizes the country’s social welfare.

If $e^* < \min\{\hat{e}, \bar{e}\}$, then for all $e_L < \min\{\hat{e}, \bar{e}\}$, country $G$ is better off when having $e_G = e_L$ and retaining all the banks than having $e_G < e_L$ and letting all bankers register in $L$. Since country $L$ strictly prefers $e_L > e_G$ as discussed above, the two countries race to establish better supervision to attract banks. Country $G$ “wins” by setting $e_G = \min\{\hat{e}, \bar{e}\} > e^*$; its welfare is lower than in the absence of country $L$ due to over supervision, nonetheless.

We have the following proposition:

**Proposition 1.7.2.** In a two-country setting without bailout guarantees, there is a regression toward over supervision if $e^* < \min\{\hat{e}, \bar{e}\}$, for $e^*$, $\hat{e}$ and $\bar{e}$ given in (1.14), (1.25) and (1.27), respectively. Social welfare of the country with a banking sector is reduced as a result.

**Proof.** See discussion above \[\square\]
1.7.2.2 Bailouts and Regulatory Shopping

Let us now consider the case with guaranteed bailouts in both countries, but there is no supervision. I do not consider supervision in this case because effective supervision cannot be motivated, as argued in Section 1.6.2.3.

Recall that in a bailout-guaranteed environment, depositors are willing to lend unlimitedly at 0 interest rate. Bankers’ utility, therefore, strictly decreases with the capital requirement they have to comply with. It follows that neither country in this case has any incentive to set their requirement higher than $k^0$, for $k^0$ given in (1.16) in Section 1.6.2.1. Furthermore, for capital requirements below $k^0$, bankers will pursue risky projects.

Given country $G$’s capital requirement $k_G \leq k^0$, country $L$’s welfare is given as follow:

$$W_L(k_L) = \begin{cases} w + \alpha b U_r^b(k_L) - \beta b(1 - r)\tau(k_L) & \text{if } k_L < k_G, \\ w & \text{otherwise,} \end{cases}$$

(1.28)

for $U_r^b(k_L)$ is a banker’s utility facing capital requirement $k_L$ and pursuing risky projects, given by:

$$U_r^b(k_L) = 1 - k_L + \frac{r}{\delta}[R - (1 - k_L)],$$

and $\tau(k_L)$ is the cost of bailing out one failed bank, given by:

$$\tau(k_L) = 1 - k_L - \xi.$$
who register with country $L$ are risky. In other words, if a banker is to take on a safe project, he may as well remain in country $G$.

Let us denote:

$$F(k_L; \alpha, \beta) = \alpha b U_r^b(k_L) - \beta b(1 - r) \tau(k_L)$$

$$= \left[ \beta(1 - r) - \alpha(1 - \frac{r}{\delta}) \right] k_L + \alpha \left[ 1 + \frac{r}{\delta}(R - 1) \right] - \beta(1 - r)(1 - \varepsilon)$$

Intuitively, $F(k_L; \alpha, \beta)$ is the utility country $L$ can extract from housing the headquarters of foreign banks. When maximizing its welfare, country $L$ will accordingly set $k_L$ to maximize $F(k_L; \alpha, \beta)$. Thus, a necessary condition for $k_L$ is $F(k_L; \alpha, \beta) \geq 0$.

Assuming $\alpha(1 - \frac{r}{\delta}) \neq \beta(1 - r)$, let:

$$\hat{k} = -\frac{\alpha \left[ 1 + \frac{r}{\delta}(R - 1) \right] - \beta(1 - r)(1 - \varepsilon)}{\alpha(1 - \frac{r}{\delta}) - \beta(1 - r)}$$

then $F(\hat{k}; \alpha, \beta) = 0$.

Furthermore, $\hat{k} = k^0$ when

$$\alpha = \beta \cdot \frac{(1 - r)(\frac{\Delta P}{\Delta r} - \varepsilon)}{\frac{\Delta P}{\Delta r} + \frac{r}{\delta}(R - \frac{\Delta P}{\Delta r}) \cdot \Omega}.$$  

It is straightforward to show that:

$$0 < \Omega < \frac{1 - r}{1 - \frac{r}{\delta}} < 1.$$  

We then have the following proposition:

**Proposition 1.7.3.** *International regulatory requirements are determined as follows:*
• For $\beta \frac{1-r}{1-\delta} \leq \alpha \leq 1$, $k_G = k_L = 0$. All bankers remain in country $G$ and pursue risky projects.

• For $\beta \Omega < \alpha < \beta \frac{1-r}{1-\delta}$, $k_G = 0$ and $k^0_L > k_L \geq \max\{\bar{k}, 0\}$. All bankers register in country $G$ and pursue risky projects.

• For $0 \leq \alpha \leq \beta \Omega$, country $G$ sets $k_G = k^0$ while country $L$ sets $k_L \geq k^0$. All bankers remain in country $G$ and pursue safe projects.

Proof. See Appendix A.3

The intuition of Proposition 1.7.3 is as follows. The international capital requirements depend on two factors: the degree ($\alpha$) to which the small country (country $L$) benefits from housing headquarters of foreign banks and the share of the burden ($\beta$) it has to bear in case of banking failures. The larger the $\alpha$, the more aggressively the small country wants to attract foreign banks. Respectively, the smaller the $\beta$, the more enthusiastic the small country is in welcoming risky foreign banks. Hence, for sufficiently large $\alpha$ and small $\beta$, the small country sets a low capital requirement that leads to risky banking. The large country (country $G$) is then induced to do the same.

1.8 Conclusion

The paper sets out to study the impacts of supervisory shopping in the banking sector. To do so, the paper employs a three-tier principal-agent framework of banking supervision and analyze the framework in alternative settings. The paper shows that supervisory shopping can inspire a “race to the top” among the supervisors, and hence, lead to a safe banking environment with a high level of
social welfare. However, the result is contingent on the principal’s commitment to not bail out distressed banks. When that commitment is not credible, the paper argues that it is difficult to motivate efficient supervision. Supervisors enter a “race to the bottom” to attract banks as banks look for weak supervisors that enable them to take more risks. As a consequence, the principal has to rely entirely on capital regulation to bring about a safe banking environment.

Furthermore, the paper extends the analysis to consider the issuance of contingent convertible bonds (CoCos) in place of deposits and also supervisory shopping in an international context. The paper finds that CoCos are socially preferable to deposits as they can absorb losses in case of project failures. Meanwhile, supervisory shopping in an international context may lead to over-supervision or under-regulation of the banking sector. As discussed in previous sections, these results are intuitive and comparable to many instances observed in reality.

The current framework in this paper takes a parsimonious approach and makes a number of simplifying assumptions, including the structure of investment projects, of supervisory contracts and of CoCos. While these assumptions do not affect the qualitative results presented in this paper, the variations of these structure can provide more insights into the nature of supervisory shopping. For instance, the riskiness of investment projects can be correlated and dependent on the number of banks picking a particular project. That would create room for “strategic complementarity” as discussed by Farhi and Tirole (2012). The effect of such changes on supervisory competition and on welfare is left for future study.
Chapter 2

Too Big To Fail: Toward an Optimal Regulation

This chapter is joint work with Chang Ma at the Johns Hopkins University.

2.1 Introduction

During the recent financial crisis (2007-09), government interventions in failing financial institutions, such as American International Group (AIG), fueled a resurgence of interest in the notion of “Too Big To Fail” (TBTF). The notion refers to the idea that some financial institutions are so large and complex that their failures are costlier to society than the required expenses to save them. The idea began to attract public attention in the 1970s when the Federal Deposit Insurance Corporation (FDIC) repeatedly bailed out institutions that it deemed “essential” to the community. In 1991, TBTF received an official status in the financial industry when Congress passed the FDIC Improvement Act, authorizing the FDIC to grant special treatments to a number of large banks.

\footnote{For a discussion on FDIC’s open bank assistance, see \textit{Federal Deposit Insurance Corporation} (1998)}
Although Congress attempted to abolish the notion of TBTF with the passage of the Dodd-Frank Act in 2010, the Act created a new category—Systemically Important Financial Institutions (SIFIs)—which not only included all previously TBTF institutions but also many new ones. With TBTF continuing to exist, how should it be regulated?

In this paper, we provide a simple framework of banking and examine a range of possible TBTF policies, including a minimum Tier-1 or Tier-2 capital requirement, tax on size, and cap on size (either via a direct cap or a maximum leverage ratio requirement). In the model, we consider a size-dependent cost of bank resolution, which provides the standard justification for providing assistance to open banks, as well as a cost of raising public funds to provide such assistance. We find that each policy helps to improve social welfare in comparison with its absence; however, to reach the first-best level of social welfare, it is necessary to impose some form of size restriction on the banker. This is because, under all other policies, our representative banker who has access to cheap deposits also has a strong incentive to expand her investment scale beyond the socially optimal level.

The issues surrounding TBTF have several layers. First, quite literally, banks can be too big to fail. This is because bank failure is typically costly to creditors and depositors, as well as disruptive to the local and even national economies (Bernanke, 1983; Chabot, 2011; Bernanke, 2013). The larger the bank, the more costly and disruptive its failure will be (White and Yorulmazer, 2014; McAndrews, Morgan, Santos, and Yorulmazer, 2014). When a large bank finds itself on the brink of collapse, the government is inclined to intervene in the form of recapitalization by using public funds (i.e., a bailout). Second, knowing
that the government will intervene, banks have a strong incentive to become TBTF. Naturally, a bank that has received either the implicit or explicit status of TBTF will face less scrutiny from the market and will be able to raise more and cheaper debts (Jacewitz and Pogach, 2011; Strahan, 2013; Santos, 2014). Furthermore, TBTF banks will be more willing to gamble with their investments (Davila, 2012; Afonso, Santos, and Traina, 2014; Gropp, Gründl, and Guettler, 2014). Third, on anticipating such intervention and the banks’ behavior, authorities have tried to regulate those banks that are (or may become) TBTF. For instance, under the authority of the FDIC Improvement Act, banks that received the TBTF status (implicitly or explicitly) were subject to a broader scope of regulation and supervision. However, TBTF banks continue to get larger in good times and require ever more public assistance in bad times (Strahan, 2013).

Regulating TBTF is not a simple task. According to Stern and Feldman (2004), it is difficult to identify and measure the TBTF problem because financial markets have grown not only in size but also in complexity. Furthermore, the benefit that TBTF institutions receive is mostly at the margin, which can vary greatly across firms of different sizes with different portfolio compositions and performance histories (Ennis and Malek, 2005). Even considering the negative impacts of TBTF as given, optimal regulation remains debatable. For instance, Johnson and Kwak (2011) argued for a straightforward cap on size and called for division of the largest financial institutions in the United States. Others, however, strongly resisted the idea for fear of inhibiting innovation and economies of scales (Krugman, 2010, April 1, for instance, see). Indeed, a number of studies have found evidence of economies of scale in banking (Hughes
and Mester, 1998; Feng and Serletis, 2010; Wheelock and Wilson, 2012; Kovner, Vickery, and Zhou, 2014).

Furthermore, another fundamental concern regarding TBTF is the use of public funds to assist open banks. Strahan (2013) provides an excellent survey of the issue, arguing that TBTF is partly due to—and always reinforced by—the government’s commitment to assist large financial institutions in distress. The justification for an ex-post intervention can be traced back to Bagehot (1873), who explained the need and presented the principles for lending of last resort, and to Diamond and Dybvig (1983), whose model provides the rationale for policy actions that prevent widespread contagion of liquidity shocks. A formal argument for bailouts can also be found in the representation hypothesis by Dewatripont and Tirole (1994), who argue that depositors are too small, and thus need protection. Based on these reasons, governments have repeatedly provided bailouts to failing institutions throughout history, and they seemingly use more of the taxpayers’ money each time. To address this issue, a growing body of literature has advocated for “bail-in” regulation, thereby shifting the burden of saving failing banks from taxpayers to holders of high-yielding bonds. In particular, Sommer (2014) and Flannery (2014) support the use of convertible debts at the largest financial institutions as a counter-measure to the moral hazard of TBTF.

Building on previous findings, we provide a two-period model of banking that focuses on the notion of TBTF. In our model, during the first period, a representative banker who bears a one-time fixed cost of investment borrows to invest in a risky project. If the banker cannot repay her debt in the second period, she declares bankruptcy and liquidates her assets at an increasing and
convex cost. As the liquidation cost is a deadweight loss to society, Congress, which is a welfare-maximizing entity, is tempted to bail out the banker via taxation of depositors. While the model considers a fixed marginal cost of taxation, Congress strives to strike a balance between the social benefits and costs of a bailout. Anticipating the bailout, however, the banker excessively increases the size of her investment and worsens the taxpayers’ burden. To manage the bailout cost, Congress can impose a minimum capital requirement, size cap, or size tax. Then, we show that the effectiveness of a minimum capital requirement is limited by the banker’s available equity. Meanwhile, a cap on size, while helping to reduce the cost to taxpayers, may inhibit the banker’s economies of scale. From a welfare perspective, a tax on size is equivalent to a cap on size. To address the government’s bailout incentive, we consider the use of contingent convertible bonds (CoCos) as an additional layer of regulatory capital. Not only do CoCos eliminate the use of taxpayer money, but they also help to reach the first-best welfare level.

It is noteworthy that, in this paper, we aim to highlight the discussion about TBTF in terms of bank size, and thus choose to simplify our main model with respect to systemic risk and its associated regulations. As Afonso, Santos, and Traina (2014), Cetorelli, McAndrews, and Traina (2014), and Laeven, Ratnovski, and Tong (2014) argue, banks tend to become larger, riskier, and more complex simultaneously. Their complexity can generate systemic risk; in other words, the failure of one institution can lead to a wave of asset fire sales and credit flow disruptions in the financial system, such as the case of Lehman Brothers. For this reason, researchers and policymakers have spent a great deal of effort understanding, measuring, and mitigating systemic risk, especially over
the last few years. However, bank size regulation remains a crucial aspect of TBTF. With the failure of Lehman Brothers, for instance, the resolution, which happened well after the financial crisis had passed, recovered less than 30 cents on the dollar for creditors, at a cost of more than $9 billion in administrative and other expenses (Fleming and Sarkar, 2014). Meanwhile, Brewer III and Jagtiani (2013) examined the data on bank mergers in the United States between 1991 and 2004, finding that banks were willing to pay additional premiums in acquisitions that expanded their size into the TBTF regime. As size continues to play an important role in bankers’ business decisions and policymakers’ responses, the pros and cons of bank size regulation require more attention.

For the sake of completeness, we consider an extension of the main model in Section 2.7. The extension includes multiple bankers and their choices in terms of the investment scale and the correlation of their returns. Similar to the findings in the aforementioned papers, we also find that bankers have a strong incentive to concentrate their risk and increase their correlations, while the per-bank cost of bailouts increases at the expense of taxpayers. However, our policy proposal, which combines size regulations and CoCos, is robust to the extension, because the policy combination forces banks to raise enough CoCos to eliminate both liquidation and bailout outcomes.

The paper proceeds as follows. Section 2.2 reviews the relevant literature. Section 2.3 describes the model and examines the market outcome in a laissez-faire environment. Section 2.4 provides a rationale for the government’s ex-post intervention as well as highlighting the moral hazard of a bailout. Section 2.5 considers size regulation. Section 2.6 proposes an optimal policy combination

\(^3\)For an exposition, see Acharya, Cooley, Richardson, and Walter (2010).
that can help to achieve the first-best level of social welfare. Section 2.7 extends the model to consider systemic risk. Section 2.8 concludes.

2.2 Literature Review

There are two bodies of literature that are closely related to this paper. The first one, which examines the interactions between bank size and the decisions of bankers and central bankers, provides rationales for TBTF and offers a range of optimal bailout strategies to offset it. The second one examines alternative ex-ante prudential policies to prevent banks from becoming TBTF. However, this latter strand of literature, while placing attention on systemic risk regulations, has not made conclusive statements regarding bank size regulation. The two bodies of literature are discussed below.

The literature regarding the rationales behind TBTF and bailout policies is expansive. Freixas (1999) provides a game-theoretic model of bank failure under the crucial assumption that the cost of bank liquidation increases with bank size. If the regulator cannot commit to an ex-ante policy, the model results in the regulator’s pure strategy of bailing out sufficiently large banks. In this case, banks anticipate the policy and structure their liabilities to maximize the value of a bailout. If the regulator can commit to an ex-ante policy, however, it is optimal for the regulator to follow a mixed strategy. In this case, banks are qualified for—but not guaranteed—a bailout when their uninsured funding is sufficiently small. Similarly, Goodhart and Huang (1999) analyze TBTF from the perspective of the lender of last resort (LOLR). In their model, the LOLR faces a request for liquidity injection from an illiquid bank, which may or may
not be insolvent. If the LOLR complies and the bank is actually insolvent, there is a direct loss of capital. If the illiquid bank’s request is declined, however, the resulting failure triggers a loss of confidence by the bank’s depositors that is socially costly. Given that the cost of bank failure rises (with respect to bank size) faster than the cost of bank rescue, the authors find that a sufficiently large bank will always receive liquidity injections.

The link between bank size and ex-post bailouts has also been a subject of contention in this literature. Acharya and Yorulmazer (2007) provide a model that demonstrates a herding behavior by small banks. In their analysis, when the number of failures is small, the regulator will let other banks acquire the failed banks. However, when the number is large, the social cost is sufficiently high that an ex-post bailout is optimal. Anticipating the regulator’s ex-post decision, small banks tend to correlate their risk of failure to increase their bailout subsidy. The authors then argue that, in this case, a TBTF bank will differentiate itself from the small banks because its bailout subsidy does not increase with the herd. Meanwhile, a recent paper by Davila (2012) contends that the presence of large banks exacerbates the risk-taking behavior of small banks and can lead to higher bailout costs. Using a model with a continuum of small banks and a definitive number of large banks, the author shows that by internalizing their size, large banks take on more risk to increase their chances of receiving bailouts. Even though small banks cannot directly influence the equilibrium bailout probability via their individual leverage decisions, they handle more risk in the presence of large banks, and hence, in aggregate, increase the probability of bailouts. Size, as Davila concludes, does matter.

With respect to this literature, our paper focuses on ex-ante regulations
rather than bailout policies, while employing similar assumptions on the costs of bank failure and rescue. In particular, we focus on bank size regulation, and demonstrate that TBTF can be resolved with a combination of policies, including a cap on size and the use of CoCos.

While a general discussion of banking regulation can be found in Dewatripont and Tirole (1994) and Freixas and Rochet (1997), a number of recent papers have considered specific ex-ante prudential regulations of TBTF. Employing similar modeling techniques as Acharya and Yorulmazer (2007), Acharya (2009) finds that banks tend to invest in the same industry, increase the correlation of their returns and failures, and thereby extract greater bailout subsidies. The paper then proposes a capital requirement that considers the banks’ joint risk, and shows that such a policy can alleviate systemic risk. Similarly, Farhi and Tirole (2012) examine a model in which banks tend to hold little liquidity and take on too much correlated risk. Therefore, in their model, it is optimal for the regulator to impose an ex-ante liquidity requirement to eliminate ex-post bailout equilibriums. Furthermore, Calomiris and Herring (2013), Chen, Glasserman, Nouri, and Pelger (2013), and Sommer (2014) propose the use of convertible debts to ensure bank solvency. Directly tackling the banks’ incentive to take risk, Acharya, Pedersen, Philippon, and Richardson (2010b) build on their earlier framework in Acharya, Pedersen, Philippon, and Richardson (2010a) and argue that the optimal policy to regulate systemic risk is a “Pigouvian tax.” This tax, which is based on the banks’ expected losses in a systemic crisis, varies with the banks’ size, leverage, risk, and correlation with the rest of the financial sector.
Our paper differs from this body of literature in considering the interaction between bank size and regulations. While focusing on mitigating systemic risk, the aforementioned papers do not address the possible underinvestment problem. In our paper, we focus on this issue and show that although banks tend to become TBTF in the absence of regulations, some ex-ante policies may inefficiently restrict the banks’ economies of scale.

2.3 The Model

Let us consider a two-period economy, \( t = 1, 2 \), with a continuum of depositors of mass 1 and a representative entrepreneurial banker, all of whom are risk-neutral.

Depositors (collectively) receive an arbitrarily large endowment, \( w \), at \( t = 1 \) and have rational expectations and preferences in a linear consumption function:

\[
U^d = E_1[c^d_1 + c^d_2],
\]

where \( c^d_t \) is their time-t consumption.

The banker receives an endowment, \( e_0 \), at \( t = 1 \), and has rational expectations and preferences in a linear consumption function:

\[
U^b = E_1[c^b_1 + c^b_2],
\]

where \( c^b_t \) is her time-t consumption.

We assume that the banker has access to an investment technology that is not available to depositors. To initiate the technology, a fixed cost \( f < e_0 \) is needed. At \( t = 1 \), the banker chooses an investment scale \( I \geq 0 \). The investment
yields $\rho(I)I$ with probability $\alpha$, and $\rho I$ with probability $1 - \alpha$. We assume that $\rho < 1$ and $\rho(I)$ is a real-valued function of $I$.

Furthermore, we assume that $\rho(I)I$ is strictly increasing and concave. Intuitively, while the payoff in the good state increases with the investment scale, the project has a strictly diminishing marginal return.

To invest, the banker can use $E$ units of her own equity, and raise additional $D = I + f - E$ units of deposits at $t = 1$. Let $R = 1 + r$ denote the return on deposits that the banker promises to depositors. We assume that the banker has limited liability. If she cannot meet her debt obligation at $t = 2$, she can declare bankruptcy.

In case of a default, the bank’s assets are liquidated at $\lambda(I)$ per unit of asset. We assume that the bigger the investment project, the more costly the liquidation process; however, the total amount of recovered assets in liquidation, $\lambda(I)I$, will increase with the project size. Formally, we have: $\lambda'(I) < 0$, $[\lambda(I)I]' > 0$, and $[\lambda(I)I]'' < 0$, for $I \geq 0$. Furthermore, $\lambda(0) = 1$ and $\lim_{t \to w}[\lambda(I)I]' = 0$.

Let $I_0 = \frac{e_0 - f}{1 - \rho}$. Then, $I_0$ is the maximum amount of investment the banker can have when issuing only risk-free deposits.

**Assumption 1. Marginal Return of the Project**

$$\alpha[\rho(I_0)I_0]' + (1 - \alpha)\rho > 1$$

Assumption 1 states that the project has a high marginal return at $I_0$.

**Assumption 2. Social Value of the Project**

There exists $I_p > I_0$, such that:

$$\alpha \rho(I_p)I_p + (1 - \alpha)\rho \lambda(I_p)I_p > I_p + f.$$
Assumption 2 ensures that the project has a positive social value for a sufficiently large investment, even when accounting for the cost of liquidation.

Note that the social value of the investment project at \( I = I_0 \) cannot be inferred from Assumptions 1 and 2. In fact, we impose Assumption 3 below to make sure that the project is not socially desirable at \( I = I_0 \). This assumption simplifies the analysis (without changing the main results) by eliminating the case in which the banker issues only risk-free deposits.

**Assumption 3.**

\[
I_0 + f \geq \alpha \rho(I_0)I_0 + (1 - \alpha)\rho I_0.
\]

Finally, we assume that the return in the good state is bounded above by some large value \( M \). The upper bound can be interpreted as a capacity constraint, induced by either technology or equipment.

**Assumption 4. Limited Capacity**

\[
\exists M > 0 \text{ such that } \rho(I)I < M, \forall I \geq 0.
\]

In the laissez-faire environment, the banker proposes \( E, I, \) and \( r \) to maximize her utility. To invest, she has three options: (1) use only her own money, (2) issue risk-free deposits, or (3) issue risky deposits to investors. As noted above, Assumption 3 implies Options 1 and 2 are not profitable.

We now proceed with the analysis of Option 3.

**Risky Deposits:**

As investing (up to) \( I = I_0 \) is not profitable by Assumption 3, the banker must risk liquidation and consider an investment scale, where \( I > I_0 \). Hence,
assuming $I > I_0$, the banker chooses $E$, $I$, and $r$ that solve:

$$\max_{E, I, r} U^b(E, I, r) = e_0 - E + \alpha \left( \rho(I)I - (1 + r)(I + f - E) \right)$$

s.t. $e_0 \geq E$ \hfill (2.1)

$$\rho(I)I \geq (1 + r)(I + f - E)$$ \hfill (2.2)

$$\alpha(1 + r)(I + f - E) + (1 - \alpha)\rho\lambda(I)I \geq I + f - E$$ \hfill (2.3)

Condition (2.2) limits the amount the banker can credibly promise to pay back. Condition (2.3) is the depositors’ Participation Constraint, ensuring that depositors can at least break even.

Note that because we are considering $I > I_0$, Condition (2.3) implies that $r > 0$.

The Lagrangian and the relevant Kuhn–Tucker conditions are as follows:

$$\mathcal{L} = e_0 - E + \left( \alpha + \gamma_2 \right) \left( \rho(I)I - (1 + r)(I + f - E) \right) + \gamma_1 (e_0 - E) + \gamma_3 \left( \alpha(1 + r)(I + f - E) + (1 - \alpha)\rho\lambda(I)I - I - f + E \right)$$

$$\frac{\partial \mathcal{L}}{\partial E} = -1 + \alpha(1 + r) - \gamma_1 + \gamma_2(1 + r) + \gamma_3(1 - \alpha(1 + r)) \leq 0$$

$$\frac{\partial \mathcal{L}}{\partial I} = \left( \gamma_2 + \alpha \left( [\rho(I)I]' - (1 + r) \right) + \gamma_3 \left( (1 - \alpha)\rho\lambda(I)I' - 1 + \alpha(1 + r) \right) \right) = 0$$

$$\frac{\partial \mathcal{L}}{\partial r} = \left( - \alpha - \gamma_2 + \gamma_3 \alpha \right)(I + f - E) = 0$$

Claim 1. Depositors’ participation constraint is always binding: $\gamma_3 \neq 0$.

Proof. Suppose that $\gamma_3 = 0$, then $\frac{\partial \mathcal{L}}{\partial r} < 0$. Contradiction. \hfill \square
Given Claim 1 and \( I > I_0 > e_0 - f \), the Kuhn–Tucker conditions above become:

\[
-(1 + \gamma_1) + \gamma_3 \leq 0, \tag{2.4}
\]

\[
\alpha\left[\rho(I)I\right]' + (1 - \alpha)\left[\lambda(I)I\right]' = 1, \tag{2.5}
\]

\[-(\alpha + \gamma_2) + \alpha \gamma_3 = 0. \tag{2.6}\]

Let \( B(I) = \alpha \rho(I)I + (1 - \alpha)\lambda(I)I - f, \) then \( B(I) - I \) is the total expected profit from an investment \( I. \) As \( \rho(I)I \) and \( \lambda(I)I \) are strictly concave, \( B(I) - I \) is strictly concave. Additionally, \([\rho(I)I]'\) is decreasing in \( I. \) Together with Assumption 2, \( B(I) - I \) is therefore positive and increasing at \( I = I_0. \)

**Lemma 2.3.1.** There exist \( I_1 \) and \( I_2 \) such that \( I_1 < I_p < I_2 \) and \( B(I_1) - I_1 = B(I_2) - I_2 = 0. \)

**Proof.** By Assumptions 2 and 3, \( B(I_0) - I_0 \leq 0 < B(I_p) - I_p. \) As \( B(I) - I \) is continuous, there must exist \( I_1 < I_p, \) such that \( B(I_1) - I_1 = 0. \)

By Assumption 4, we have \( 0 = \alpha(M - M) > \alpha(\rho(M)M - M) > B(M) - M. \)
Hence, \( B(I) - I \) is negative at \( I = M. \) Furthermore, \( \alpha M > \alpha \rho(I_p)I_p > I_p \) by Assumptions 2 and 4. It follows that \( M > I_p. \)

Thus, by continuity, there must exist \( I_2 \in (I_p, M) \) such that \( B(I_2) - I_2 = 0. \)

We now have the following proposition.

**Proposition 2.3.2.** In the laissez-faire environment, the banker chooses to issue risky debt to satisfy the investment scale \( I = I^{LF}, \) such that \( I_1 < I^{LF} < I_2. \)
Furthermore, the banker’s profit is strictly positive at \( I = I^{LF}. \)
Proof. First, an optimal investment scale $I$ that solves the banker’s maximization problem must satisfy Condition (2.5).

Second, note that $B(I_p) - I_p > 0$ and $B(I_1) - I_1 = B(I_2) - I_2 = 0$. As $B(I) - I$ is strictly concave, $I^{LF} \in (I_1, I_2)$ must uniquely exist, such that:

$$[B(I^{LF}) - I^{LF}]' = 0,$$

$$\Rightarrow \alpha[\rho(I^{LF})I^{LF}]' + (1 - \alpha)[\lambda(I^{LF})I^{LF}]' = 1$$

As $I^{LF}$ indeed maximizes $B(I) - I$, it also follows that $B(I^{LF}) - I^{LF} > 0$.

Now, consider the banker’s maximization problem and the Kuhn–Tucker conditions above. Recall that $\gamma_3 > 0$ by Claim 1.

If $\gamma_2 \neq 0$, then by Conditions (2.2) and (2.3), we obtain

$$\alpha \rho(I^{LF})I^{LF} + (1 - \alpha) \lambda(I^{LF})I^{LF} = I^{LF} + f - E,$$

which is a contradiction because $B(I^{LF}) - I^{LF} > 0$ and $E \geq 0$.

Thus, $\gamma_2 = 0$. It then follows that, for any $E \in [0, e_0]$, there exists $r^E$ such that $(E, I^{LF}, r^E)$ solves the banker’s maximization problem. Furthermore, it is straightforward to verify that the banker’s profit for any $E \in [0, e_0]$ is $B(I^{LF}) - I^{LF} > 0$.

The banker’s utility is:

$$U^b(E, I^{LF}, r^E) = e_0 + B(I^{LF}) - I^{LF},$$

and the expected social welfare in this case is as follows:

$$W^{LF} = w + e_0 + B(I^{LF}) - I^{LF}. $$

66
Note that in this case, we do not observe a clear role for the use of equity. This is intentional. As we focus on the banker’s eventual incentive to expand her investment scale, the model considers neither the banker’s risk shifting behavior nor her cost of equity. In the next sections, the role of equity as a regulatory tool will become apparent.

2.4 Moral Hazard of Government Bailouts

In the previous section, we considered the laissez-faire environment in which the banker borrows from depositors to invest. As the banker makes a profit from using depositors’ money, depositors expect and are completely compensated for the risk of default, including the cost of liquidation.

In this section, we consider the presence of a welfare-maximizing government entity, referred to as Congress, which has the power to tax depositors and the banker, impose banking regulations and make transfers.

If Congress chooses to tax depositors, we assume that there is a fixed marginal cost in raising public funds, \(0 < \eta < 1\).

In the event of a bank failure, Congress may choose to recapitalize (i.e., bail out) the bank. The justification for this action is to prevent the bank from declaring bankruptcy; hence, saving the public from the cost of liquidation. It follows that Congress would choose to bail out the bank if the total cost of liquidation exceeds the total cost of raising the necessary public funds to do so. Formally, the condition for a bailout is:

\[
\eta(D(1 + r) - \rho I) \leq (1 - \lambda(I))\rho I,
\]

where \(D\) is the total deposits and \(r\) is the interest rate on deposits.
We define a bailout as a transfer at $t = 2$ to the bank in the case of project failure. The transfer is financed by an ex-post tax on all depositors. If the bank receives a bailout, Congress acts as the bank’s conservator and takes control of the bank’s assets. This assumption prevents the bank from strategically declaring bankruptcy.

Furthermore, we impose the following assumption.

**Assumption 5.**

$$\frac{(1 - \lambda(I^{LF}))\rho}{1 - \rho + f/I^{LF}} > \eta.$$  

Assumption 5 implies that the marginal cost of raising public funds is sufficiently small in comparison with the marginal rate of the project return in the bad state.

Let $G(E, I) = (1 - \lambda(I))\rho I - \eta(I + f - E - \rho I)$. Then, conditional on a zero interest rate on deposits, Congress bails out the bank if and only if $G(E, I) \geq 0$.

Note that, $\frac{\partial}{\partial I} G(E, 0) = -\eta(1 - \rho) < 0$; and $\frac{\partial^2}{\partial^2 I} G(E, I) = -\rho[\lambda(I)I]' > 0$.

Given Assumption 5 and the negativity of $\lambda'(I)$, we have:

$$\frac{\partial}{\partial I} G(E, I) = \rho - \rho[\lambda(I)I]' - \eta + \eta \rho$$

$$> \rho - \rho \lambda(I) - \eta + \eta \rho$$

$$> 0,$$

for $I > I^{LF}$.

Intuitively, $G(E, I)$ is the function governing the condition for an ex-post bailout when the interest rate on deposits is zero. It is strictly convex in $I$, and its first derivative with respect to $I$ is negative for small $I$ and becomes positive.
for sufficiently large $I$. It follows that there must exist $I_{G}^{Min} \in (0, I^{LF})$, such that $\frac{\partial}{\partial I} G(E, I_{G}^{Min}) = 0$.

In this section, we analyze market outcomes in the absence of regulation, and under a minimum equity capital requirement. We find that the banker consistently expands her investment scale to raise her expected profit at the expense of taxpayers.

2.4.1 Too Big To Fail in the Absence of Regulation

Suppose at $t = 1$ the government does not impose any regulation on the bank.

First, the banker is able to raise deposits as if the bailout is guaranteed. This is because if depositors and the banker do not expect an ex-post bailout, the optimal investment will be $I^{LF}$, as discussed in Section 2.3. By Assumption 5, the no-bailout expectation is time inconsistent.

Second, as depositors and the banker fully expect a bailout in the event of bank failure, the banker does not have to pay any interest on deposits. Depositors know that the bank will be bailed out and are willing to accept a zero interest rate. The banker’s equity contribution and investment scale must then satisfy the bailout condition, $G(E, I) \geq 0$.

To borrow, the banker proposes $(E, I)$ that solves:

$$\max_{E,I} U^b(E, I) = e_0 - E + \alpha \left( \rho(I) I - (I + f - E) \right)$$

s.t. $e_0 \geq E$  

$$\rho(I) I \geq I + f - E$$

$$G(E, I) \geq 0$$
The relevant Kuhn–Tucker conditions are as follows:

\[ \mathcal{L} = e_0 - E + \left( \alpha + \gamma_2 \right) \left( \rho(I)I - I - f + E \right) + \gamma_1 \left( e_0 - E \right) + \gamma_3 G(E, I) \]

\[ \frac{\partial \mathcal{L}}{\partial E} = -1 + \alpha + \gamma_2 - \gamma_1 + \gamma_3 \eta \leq 0 \] (2.11)

\[ \frac{\partial \mathcal{L}}{\partial I} = \left( \alpha + \gamma_2 \right) \left( [\rho(I)I]' - 1 \right) + \gamma_3 \frac{\partial}{\partial I} G(E, I) = 0 \] (2.12)

We specifically focus on the case in which the profit constraint (2.9) holds strict inequality, so that the banker makes a positive profit. It follows that \( \gamma_2 = 0 \).

Let \( I^\rho \) be such that \( [\rho(I^\rho)I^\rho]' = 1 \). We have the following proposition.

**Proposition 2.4.1. Bigger and Badder**

In the absence of any regulation, the TBTF banker maximizes her utility at \((E, I) = (0, I^\rho)\). The investment scale in this case is larger than in Section 2.3; that is, \( I^\rho > I^{LF} \). The banker’s utility is also higher; that is, \( U^b(0, I^\rho) > U^b(E, I^{LF}, r_E) \).

**Proof.** It is straightforward to verify that \((E, I) = (0, I^\rho)\) satisfies the Kuhn–Tucker conditions (2.11) and (2.12), where \( \gamma_1 = \gamma_2 = \gamma_3 = 0 \).

Recall that \( I^{LF} \) satisfies:

\[ \alpha [\rho(I^{LF})I^{LF}]' + (1 - \alpha) \rho [\lambda(I^{LF})I^{LF}]' = 1 \]

\[ [\rho(I^\rho)I^\rho]' = 1 + \frac{(1 - \alpha)}{\alpha} \left( 1 - \rho [\lambda(I^{LF})I^{LF}]' \right) \]

\[ > [\rho(I^\rho)I^\rho]' \].

70
As $[\rho(I)I]'$ is strictly decreasing, $I^{LF} < I^\rho$.

As $I^\rho > I^{LF}$, it also follows that $G(0, I^\rho) > G(0, I^{LF}) > 0$, by Assumption 5. Thus, the bailout condition (2.10) is satisfied.

Furthermore, $I^\rho$ maximizes $\rho(I)I - I$. Hence,

$$U^b(0, I^\rho) = e_0 + \alpha\left(\rho(I^\rho)I^\rho - I^\rho - f\right)$$

$$> e_0 + \alpha\left(\rho(I^{LF})I^{LF} - I^{LF} - f\right) - (1 - \alpha)(1 - \rho\lambda(I^{LF}))I^{LF}$$

$$> U^b(E, I^{LF}, r^E).$$

Proposition 2.4.1 reveals two insights. First, by taking advantage of the bailout, the banker expands her investment scale beyond what the laissez-faire market would have allowed. Second, unlike the previous case, the banker intentionally chooses to commit no equity of her own. In Section 2.3, the banker is indifferent between equity and deposits. She fairly shares the profit from the investment with depositors, as she fully compensates depositors for the downside risk that they bear. In this case, however, the banker raises deposits at a zero interest rate, enjoys all the upside profits and leaves the downside losses to taxpayers.

The expected social welfare in this case is:

$$W(0, I^\rho) = w + U^b(0, I^\rho) - (1 - \alpha)(1 + \eta)(I^\rho + f - \rho I^\rho).$$

It is important to note that the social welfare in this case can be either higher or lower than in Section 2.3. While the presence of government intervention is socially beneficial (the bailout condition), the banker raises her investment,
potentially inflicting a large bill on taxpayers. Nonetheless, it is clear from Proposition 2.4.1 that the banker benefits from the bailout.

2.4.2 Capital Requirement

Now, suppose that Congress imposes a minimum capital requirement, $E$, at $t = 1$, such that $e_0 \geq E$.

Similar to Section 2.4.1, the banker solves the maximization problem (2.7), subject to (2.8), (2.9), and (2.10), with an additional constraint, $E \geq E$.

Solving similar Kuhn–Tucker conditions, the banker’s solution in this case is $(E, I) = (E, I^p)$.

Thus, Congress solves the following problem.

$$\max_{E \leq e_0} W(E, I^p) = w + e_0 - E + \alpha \left( \rho(I^p)I^p - (I^p + f - E) \right)$$

$$- (1 - \alpha)(1 + \eta)(I^p + f - E - \rho I^p).$$

The first-order condition is:

$$-1 + \alpha + (1 - \alpha)(1 + \lambda) = \lambda(1 - \alpha) > 0,$$

which implies that $E = e_0$.

The expected social welfare in this case becomes:

$$W(e_0, I^p) = w + U^b(e_0, I^p) - (1 - \alpha)(1 + \eta)(I^p + f - e_0 - \rho I^p)$$

We have the following Lemma.

Lemma 2.4.2. By implementing a minimum equity capital requirement, $E = e_0$, Congress raises the expected social welfare in comparison with that in the case with no regulation: $W(e_0, I^p) > W(0, I^p)$.  

72
Proof. Rewriting $W(e_0, I^\rho)$, we have:

\[
W(e_0, I^\rho) = w + e_0 + \alpha \left( \rho(I^\rho) I^\rho - I^\rho - f \right) - (1 - \alpha)(1 + \eta)(I^\rho + f - \rho I^\rho)
\]

\[+ \lambda(1 - \alpha)e_0\]

\[= W(0, I^\rho) + \lambda(1 - \alpha)e_0.\]


The intuition of Lemma 2.4.2 is simple. By requiring the banker to contribute her own equity capital, Congress is able to reduce the amount that is needed in the event of a bailout. Although a minimum equity capital requirement does not alter the banker’s investment strategy, it forces the banker to pay, if only partially, for the bailout of her bank.

Once again, because the banker’s equity capital is limited, it is unclear whether social welfare in this case is higher or lower than in the case of a free market. In the next section, we consider bank size regulation and show that social welfare can be improved. Nonetheless, bank size regulation is not without its own weaknesses.

2.5 Bank Size Regulation: Better, But Not the Best

2.5.1 A Cap On Size

Let us consider the case in which Congress implements a minimum equity capital requirement, $E$ and a direct cap on size, $I$. This combination is strong in the sense that the banker has little choice left. It is straightforward to verify that
she will contribute the minimum equity required and invest at the maximum scale allowed.

Congress, while internalizing the possibility and cost of a bank bailout, solves the following problem.

\[
\max_{E,I} W(E,I) = w + e_0 - E + \alpha \left( \rho(I)I - (I + f - E) \right) - (1 + \eta)(1 - \alpha)(I + f - E - \rho I)
\]  
\quad \text{s.t.} \quad e_0 \geq E \quad (2.13)

\[
G(E,I) \geq 0 \quad (2.14)
\]

\[
\rho(I)I > I + f - E \quad (2.16)
\]

Condition (2.16) is the banker’s Incentive Compatibility Constraint, ensuring that the profit from an investment project is positive. We consider the strict inequality of this condition because social welfare increases beyond the depositors’ and banker’s endowment only when the banker makes a profit. Note that in expectation of a bailout, depositors are willing to lend at a zero interest rate.

The relevant Kuhn–Tucker conditions for the above mentioned maximization problem can be written as follows.

\[
(1 - \alpha)\eta - \mu_1 + \mu_2 \eta \leq 0, \quad (2.17)
\]

\[
\alpha \left[ \rho(I)I' - \alpha - (1 - \alpha)(1 + \eta)(1 - \rho) \right] = -\mu_2 \frac{\partial}{\partial I} G(E,I). \quad (2.18)
\]

Let \( \bar{I} \) be such that:

\[
\left[ \rho(\bar{I})\bar{I}' \right] = 1 + \frac{(1 - \alpha)}{\alpha}(1 + \eta)(1 - \rho).
\]
Then, we have the following lemma:

**Lemma 2.5.1.** $I^\rho > \bar{I} > I^{LF}$.

**Proof.** First, we have:

\[
[\rho(\bar{I})\bar{I}]' = 1 + \frac{(1 - \alpha)}{\alpha}(1 + \eta)(1 - \rho) > 1 = [\rho(I_{\rho})I_{\rho}]'.
\]

Since $[\rho(I)I]'$ is decreasing, $I^\rho > \bar{I}$.

Second, Assumption 5 and the strict negativity of $\lambda'(I)$ imply the following:

\[
(1 + \eta)(1 - \rho) < 1 - \lambda(I^{LF})\rho < 1 - [\lambda(I^{LF})I^{LF}]'\rho.
\]

Hence, $[\rho(\bar{I})\bar{I}]' < [\rho(I^{LF})I^{LF}]'$. Since $[\rho(I)I]'$ is decreasing, $\bar{I} > I^{LF}$. 

We have the following proposition.

**Proposition 2.5.2.** The optimal allocation $(E, I)$ that solves the maximization problem (2.13), subject to (2.14), (2.15), and (2.16), is $(E, I) = (e_0, \bar{I})$. Furthermore, $W(e_0, \bar{I}) > W^{LF}$ and $W(e_0, \bar{I}) > W(e_0, I^\rho)$.

**Proof.** Consider the maximization problem (2.13), subject to (2.14) and (2.16). This problem maximizes the same objective function, but with less constraints. The Kuhn–Tucker conditions become:

\[
(1 - \alpha)\eta - \mu_1 \leq 0,
\]

\[
\alpha[\rho(I)'I] - \alpha - (1 - \alpha)(1 + \eta)(1 - \rho) = 0.
\]

The first inequality condition implies that $\mu_1 > 0$. Hence, $E = e_0$. The second condition implies that $I = \bar{I}$. 

75
Thus, if \((E, I) = (e_0, \bar{I})\) satisfies the bailout condition (2.15), \((e_0, \bar{I})\) must also be the solution to the original maximization problem with full constraints.

**Case 1:** If \(G(e_0, I_{G}^{Min}) > 0\), then \(G(e_0, \bar{I}) > 0\).

**Case 2:** If \(G(e_0, I_{G}^{Min}) \leq 0\), then there exist \(I_1\) and \(I_2\) such that \(I_1 \leq I_{G}^{Min} \leq I_2\) and \(G(e_0, I) = 0\) at \(I \in \{I_1, I_2\}\). This is because \(G(E, I)\) is a strictly convex function, which is positive at \(I = 0\), decreasing for small \(I\), and increasing for sufficiently large \(I\).

Because \(G(E, I)\) is linear and increasing in \(E\), and because of Assumption 5, we obtain the following:

\[
G(e_0, I^{LF}) > G(0, I^{LF}) > 0 = G(e_0, I_2).
\]

Since \(I_2\) and \(I^{LF}\) are greater than \(I_{G}^{Min}\) and \(G(e_0, I)\) is increasing for \(I \geq I_{G}^{Min}\), it must be the case that \(I^{LF} > I_2\).

By Lemma 2.5.1, \(\bar{I} > I^{LF}\). Hence \(G(e_0, \bar{I}) > G(e_0, I_2) = 0\).

Finally, to show \(W^{LF} < W(e_0, \bar{I})\), we rewrite \(W^{LF}\) as follows:

\[
W^{LF} = w + e_0 + \alpha \rho(I^{LF})I^{LF} + (1 - \alpha)\leq\rho(\lambda(I^{LF})I^{LF} - I^{LF} - f)
\]

\[
= w + e_0 + \alpha \left(\rho(I^{LF})I^{LF} - I^{LF} - f\right)
\]

\[
- (1 - \alpha) \left( I^{LF} + f - \leq\rho(\lambda(I^{LF})I^{LF}) \right).
\]

Using Assumption 5, we then have:

\[
W^{LF} < w + e_0 + \alpha \left(\rho(I^{LF})I^{LF} - I^{LF} - f\right)
\]

\[
- (1 - \alpha)(1 + \eta) \left[ (1 - \leq\rho)I^{LF} + f \right]
\]

\[
= W(0, I^{LF}).
\]
As \((E, I) = (0, I^{LF})\) meets all the constraints of (2.14), (2.15), and (2.16), 
\((0, I^{LF})\) is a possible solution to the maximization problem (2.13), subject to those constraints. However, as demonstrated above, the solution to the problem is \((E, I) = (e_0, \bar{I}).\)

Therefore, \(W(e_0, \bar{I}) > W(0, I^{LF}) > W^{LF}.\)

Similarly, \((E, I) = (0, I^\rho)\) is also a possible solution to the same maximization problem, subject to the same constraints. It follows that \(W(e_0, \bar{I}) > W(e_0, I^\rho).\)

Proposition 2.5.2 shows the optimal levels of equity and investment. The banker’s expected utility is as follows:

\[
U^b(e_0, \bar{I}) = \alpha\left(\rho(\bar{I})\bar{I} - \bar{I} - f + e_0\right).
\]

Under the policy combination of a minimum equity capital requirement and a cap on size, Congress achieves two objectives: (1) forcing the banker to pay for the bailout to her greatest ability, and (2) implementing an investment size that internalizes the cost of the bailout. As shown in Proposition 2.5.2, social welfare is clearly greater than in all of the previous cases.

Before proceeding, we note that Congress can implement a maximum leverage ratio, \(k := \frac{I}{e_0 - f}\) and achieve the same result. Within the context of our model, an equity-and-size policy combination and a leverage-ratio requirement are therefore identical.
2.5.2 A Tax On Size

Directly tackling the banker’s incentive, we consider a distortionary tax, $\tau$, on her investment size, $I$. There are two alternative taxation methods that Congress can implement:

- First, Congress can set a marginal tax rate, $\tau$, in the first period. The tax revenue, $\tau I$, is collected in the second period if the banker’s project succeeds. The revenue is rebated to depositors. If the project fails, Congress bails out the bank, as in the previous cases.

- Second, assuming Congress has access to a zero-return storage technology, it can set and collect $\tau I$ in the first period. If the project succeeds, the tax revenue is rebated in a lump-sum fashion to the banker. If the project fails, Congress uses the tax revenues toward bailing out the bank.

Note that in the second scenario, the banker simply borrows an additional amount, $\tau I$, from depositors, for any given investment scale, $I$, that she plans to invest. Formally, the banker borrows $D = I + f + \tau I - e_0$. In the event of bank failure, Congress uses the collected tax revenue, $\tau I$, and the banker’s asset, $\rho I$, as well as an additional amount, $I + f - e_0 - \rho I$, to settle the bank’s debt. Thus, for a given investment size, the tax revenue from the banker does not help to lower the total cost of a bailout.

In both scenarios, the tax on size merely focuses on the banker’s incentive to expand her investment. From a social welfare perspective, Congress is indifferent between the two methods.

We now proceed with the first case.
Based on the previous sections, we simplify the analysis in this section by assuming that a minimum capital, $E = e_0$, is already in place.

Suppose Congress announces at $t = 1$ a marginal tax, $\tau$, on the banker’s investment, $I$, to be collected at $t = 2$. The banker borrows $D = I + f - e_0$ in the first period and solves:

$$\max_I U^\tau(I) = \alpha \left( \rho(I)I - (I + f - e_0) - \tau I \right)$$

s.t. \hspace{1cm} \rho(I)I \geq (1 + \tau)I + f - e_0$$

$$G(e_0, I) \geq 0$$

The relevant Kuhn–Tucker conditions are as follows:

$$L = \left( \alpha + \gamma_1 \right) \left( \rho(I)I - (1 + \tau)I + e_0 \right) + \gamma_2 G(e_0, I)$$

$$\frac{\partial L}{\partial I} = \left( \alpha + \gamma_1 \right) \left( \rho(I)I' - 1 - \tau \right) + \gamma_2 \frac{\partial}{\partial I} G(e_0, I) = 0$$

Similar to the previous sections, it is straightforward to check that $I = I^\tau$ is the optimal solution where:

$$\left[ \rho(I^\tau)I^\tau \right]' = 1 + \tau.$$

Thus, Congress can directly choose $I^\tau$ to maximize the social welfare:

$$\max_{I^\tau} W(I^\tau) = w + \alpha \left( \rho(I^\tau)I^\tau - (I^\tau + f - e_0) - \tau I^\tau \right)$$

$$+ \alpha \tau I^\tau - (1 - \alpha)(1 + \eta) \left( (1 - \rho)I^\tau - e_0 + f \right).$$

The second part of the above expression is the expected net tax revenue. With probability $\alpha$, the project succeeds and the banker pays Congress $\tau I^\tau$ in
tax, which is rebated to depositors. With probability $1 - \alpha$, the project fails and Congress has to raise taxes to bail out the bank.

The simplified first-order condition is as follows:

$$[ho(I^\tau)I^\tau]' = 1 + \frac{1 - \alpha}{\alpha}(1 + \eta)(1 - \rho).$$

Hence, $I^\tau \equiv \bar{I}$.

We then have the following proposition:

**Proposition 2.5.3.** Under the tax-on-size policy, Congress maximizes social welfare by implementing a tax, $\tau$, that gives $I^\tau = \bar{I}$. As a result, $W(I^\tau) = W(e_0, \bar{I})$.

**Proof.** From the discussion above, it is clear that $I^\tau = \bar{I}$ maximizes social welfare. Furthermore,

$$W(I^\tau) = w + \alpha\left(\rho(\bar{I})\bar{I} - \bar{I} + e_0 - f\right) - (1 - \alpha)(1 + \eta)(1 - \rho)(\bar{I} - e_0 + f)$$

$$= W(e_0, \bar{I}).$$

Proposition 2.5.3 shows that a simple tax on size is equivalent to a cap on size in terms of welfare. The tax, therefore, serves as an alternative policy to a cap on size.

### 2.6 CoCos: Toward Better Regulation of Large Banks

In Sections 2.4 and 2.5, we analyzed the banker’s incentive to take advantage of a government bailout, and concluded that size regulation, in the form of either
a direct cap on size or a maximum leverage-ratio requirement, was necessary. However, size regulation is not sufficient to induce the first-best level of social welfare. This is because taxpayers’ money, which is used to foot the bill in the event of bank failure, comes at a cost.

In this section, we consider contingent convertible bonds (CoCos) as an additional layer of capital that can avoid entirely the use of taxpayer money. While the use of CoCos has been strongly advocated in the literature (Albul, Jaffee, and Tchisty, 2013; Pazarbasioğlu, Zhou, Le Leslé, and Moore, 2011; Glasserman and Nouri, 2012; Avdjiev, Kartasheva, and Bogdanova, 2013), CoCos can be a particularly important policy tool when it comes to regulating the largest financial institutions (Calomiris and Herring, 2013; Chen, Glasserman, Nouri, and Pelger, 2013; Sommer, 2014).

Similar to the previous section, we assume that a minimum Tier-1 equity capital of $e_0$ is already in place. Congress, therefore, implements a maximum leverage ratio, $k = I/(E - f)$, where $I$ is the investment size and $E$ is the Tier-2 capital on the bank’s balance sheet.

The Tier-2 capital, $E$, is defined as a sum of the Tier-1 equity capital and CoCos. Following the description of Avdjiev, Kartasheva, and Bogdanova (2013), we consider CoCos with a simple structure, as follows:

1. The banker issues CoCos in exchange for $B$ from depositors at $t = 1$,

2. The bond promises to pay $B(1 + r_B)$ to bondholders at $t = 2$,

3. At $t = 2$, if the bank is insolvent, Congress triggers a bond-to-equity conversion.

81
Note that in the absence of size-contingent regulation (such as a maximum leverage ratio), the banker still has a strong incentive to over-invest. To see this response, let us suppose that Congress implements only a Tier-2 capital requirement, \( B + e_0 \), that is independent of the investment scale. The banker raises \( B \) at the cost \( r_B \), such that \( \alpha(1 + r_B)B \geq B \), to satisfy the requirement. As the banker still has the option of raising deposits at no cost, she is free to choose \( I \) to maximize her profit, \( \alpha \left( \rho(I)I - (1 + r_B)B - (I + f - B - e_0) \right) \). The first-order condition yields \( I = I^\rho \), which is not the socially optimal level of investment.

Thus, it is necessary to incorporate size regulation to achieve the first-best allocation. As shown in Section 2.5, size cap, tax on size and a leverage requirement are equivalent in terms of welfare. In this case, we choose to consider a leverage ratio that aligns closely with the literature on CoCos.

Now, let \( I^* \) be such that:

\[
[rho(I^*)] = 1 + \frac{1 - \alpha}{\alpha} (1 - \rho).
\]

We then have the following proposition.

**Proposition 2.6.1.** The maximum leverage ratio, \( k^* = \frac{1}{1 - \rho} \), is necessary and sufficient for a no-bailout commitment. Accordingly,

- the banker chooses to invest \( I = I^* > \bar{I} \);
- the banker issues CoCos, such that \( e_0 + B = I^* (1 - \rho) + f \);
- the banker also issues risk-free deposits, \( D = \rho I^* \), at a zero interest rate;
- finally, social welfare is higher than \( W(e_0, \bar{I}) \) in Section 2.5.
Proof. See Appendix B.1

In the setting of our model, there are two types of friction that potentially reduce social welfare. One is the cost of liquidation, whereas the other is the cost of raising public funds—both are increasing with the investment size. When Congress provides an ex-post bailout, it eliminates the former; however, it imposes the latter on taxpayers. As we have shown throughout the paper, various policies help Congress reduce the burden on taxpayers and improve welfare.

Nonetheless, Proposition 2.6.1 shows that Congress can eliminate both types of frictions and achieve a higher level of welfare, especially at no cost to taxpayers. By allowing the banker to use CoCos as an additional layer of capital and by limiting leverage, Congress addresses directly the incentive of the banker. The banker, while maximizing her profit, issues only risk-free deposits, which eliminate the bankruptcy outcome and its associated cost.

2.7 Extension: Investment Concentration

In this section, we extend our benchmark model to consider the presence of two bankers, who are identical to the one described in Section 2.3. The two bankers, denoted $A$ and $B$, will choose the scale of their investment as well as the correlation of their returns.

Following the setting in Acharya and Yorulmazer (2007), we assume that there are two separate investment categories from which the bankers can choose. If the bankers choose different categories, their returns are independent. In this case, the bankers’ investment technology is exactly the same as that described in Section 2.3. The marginal costs of liquidation and bailout are also the same,
regardless if one or both banks fail.

If the bankers choose the same investment category, their returns are perfectly correlated. In other words, the bankers’ projects both succeed with probability $\alpha$ or both fail with probability $1 - \alpha$. In this case, given an investment scale, $I_i$, banker $i$’s project yields $\rho(I_i)I_i + \Delta$ in the good state and $\rho I_i$ in the bad state, for $i \in \{A, B\}$. We assume that $\Delta > 0$.

The joint distribution of bank returns and payoffs is given in Table 2.1.

<table>
<thead>
<tr>
<th></th>
<th>Bank A</th>
<th></th>
<th>Bank B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S Success; F Failure</td>
<td></td>
<td>Bank B</td>
</tr>
<tr>
<td>$i \in {A, B}$</td>
<td>S</td>
<td>F</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td>Different Categories</td>
<td>Same Category</td>
<td></td>
</tr>
<tr>
<td>Bank</td>
<td>Prob.</td>
<td>S</td>
<td>$\alpha^2$</td>
</tr>
<tr>
<td>A</td>
<td>F</td>
<td>$\alpha(1 - \alpha)$</td>
<td>$(1 - \alpha)^2$</td>
</tr>
<tr>
<td>Return</td>
<td>S</td>
<td>$\rho(I_i)I_i$</td>
<td>$\rho(I_A)I_A, \rho I_B$</td>
</tr>
<tr>
<td>F</td>
<td>$\rho I_A, \rho(I_B)I_B$</td>
<td>$\rho I_i$</td>
<td>$-$</td>
</tr>
</tbody>
</table>

Table 2.1: Banks’ payoffs and joint distribution of returns

Additionally, we assume that when banks invest in the same category and fail, there is a systemic crisis that affects the costs of liquidation and bailout. Specifically, the liquidation of bank $i$ recovers only $h\lambda(I_i)$ for $0 < h < 1$. Meanwhile, the marginal cost of raising public funds is $\eta + \theta$.

For simplicity, we assume that the extra cost, $\theta$, is sufficiently small that bailing out the banks is ex-post optimal in the relevant cases. Formally, we have the following:

**Assumption 6.**

$$\frac{\lambda(I^*)}{1 - \rho + f/I^*} \rho \geq \theta.$$
The setting in this section can be interpreted as follows. Suppose the banks have the choice to develop commercial buildings that are located either far from one another, or in the same neighborhood. If banks locate their buildings away from each other, there is no correlation. If banks invest in the same neighborhood, their returns are highly correlated. Furthermore, when both buildings are successfully built, the neighborhood becomes more attractive, which increases the value of the local real estate ($\Delta$). However, if both projects fail, the liquidation process dampens the neighborhood’s attractiveness, which further reduces the liquidation value ($0 < h < 1$). Should the government choose to bail out the banks, then it incurs an additional overhead cost, ($\theta$), to maintain both buildings.

Note that conditional on banks choosing different investment categories, all of our previous results continue to hold.

Let us now consider the laissez-faire case. As shown in Section 2.3, if the bankers choose different investment categories, they will invest $I_{LF}$ and enjoy $B(I_{LF}) - I_{LF}$ in profits.

If the bankers choose the same investment category, each bank $i$ solves the following maximization problem:

$$
\max_{E, I, r} U_b^i(E, I, r) = e_0 - E + \alpha \left( \rho(I)I + \Delta - (1 + r)(I + f - E) \right)
$$

s.t.

1. $e_0 \geq E$ \hspace{2cm} (2.20)
2. $\rho(I)I + \Delta \geq (1 + r)(I + f - E)$ \hspace{2cm} (2.21)
3. $I + f - E \leq \alpha(1 + r)(I + f - E) + (1 - \alpha)\rho h \lambda(I)I$ \hspace{2cm} (2.22)
It is straightforward to show that both bankers will choose $I_A = I_B = I_{LF}^h$ and enjoy profit $B_h(I_{LF}^h) - I_{LF}^h$, satisfying the following condition:

\[
B_h(I) = \alpha(\rho(I)I + \Delta) + (1 - \alpha)\rho h\lambda(I)I - f, \quad (2.23)
\]

\[
[B_h(I_{LF}^h) - I_{LF}^h]' = 0. \quad (2.24)
\]

It is clear that $B_h(I) - I$ is a strictly concave function, maximized at $I = I_{LF}^h$. Because $h < 1$, we have $[B(I) - I]' > [B_h(I) - I]', \forall I$. Thus, $I_{LF}^h < I_{LF}^h$.

Meanwhile, the bankers’ choice to be either independent of, or correlated with, one another depends on their expected profits. Specifically, if

\[
B(I_{LF}^h) - I_{LF}^h \geq B_h(I_{LF}^h) - I_{LF}^h, \quad (2.25)
\]

then the bankers prefer to invest in different investment categories. Otherwise, the bankers will choose the same category.

Note that inequality (2.25) holds for sufficiently small $\Delta$ and $h$. For instance, a sufficient condition for the inequality is:

\[
\alpha \Delta < (1 - \alpha)\rho h\lambda(I_h)(1 - h),
\]

which implies $B_h(I_{LF}^h) - I_{LF}^h < B(I_{LF}^h) - I_{LF}^h < B(I^{LF}) - I^{LF}$.

Let us now consider the availability of bailouts without any ex-ante regulation.

From Section 2.4.1, we know that, conditional on choosing the same investment category, bankers will contribute no capital, invest at scale $I^p$, and enjoy the benefit of a bailout.

If bankers invest in different categories, they solve a utility maximization problem that is similar to the one presented in Section 2.4.1. In particular, the
Kuhn–Tucker conditions to the bankers’ maximization problem are exactly the same as before, because the added value $\Delta$ in this case does not depend on the investment scale. It is then easy to verify that the bankers will also contribute no capital and choose the investment scale, $I^\rho$, given that the bailout condition,

$$(1 - h\lambda(I^\rho))\rho I^\rho \geq (\eta + \theta)(I^\rho + f - \rho I^\rho),$$

is satisfied.

Rewriting the bailout condition in this case, we have:

$$(1 - \lambda(I^\rho))\rho I^\rho + (1 - h)\lambda(I^\rho)\rho I^\rho \geq (\eta + \theta)(I^\rho + f - \rho I^\rho)$$

$$G(0, I^\rho) + (1 - h)\lambda(I^\rho)\rho I^\rho - \theta(I^\rho + f - \rho I^\rho) \geq 0.$$

As shown in Section 2.4.1, $G(0, I^\rho) > 0$. The latter part of the above inequality is also positive by Assumption 6. Hence, the bailout condition is satisfied.

It follows that, conditional on choosing the same investment category, the bankers will contribute no capital and expand their investment to $I^\rho$. The bankers’ utility in this case is:

$$U^b_i(0, I^\rho) = e_0 + \alpha \left(\rho(I^\rho)\rho I^\rho + \Delta - I^\rho - f\right),$$

which is greater than $U^b(0, I^\rho)$ in Section 2.4.1, for any positive $\Delta$.

Thus, given the availability of bailouts, bankers will concentrate their investment in the same category.

We have the following proposition.

**Proposition 2.7.1.** Concentration of Investment
In the laissez-faire environment, bankers choose to invest in the same category if and only if inequality (2.25) is not satisfied.

However, given the availability of bailouts, bankers concentrate their investments in the same category for any positive $\Delta$.

**Proof.** See discussion above.

Recall that Proposition 2.4.1 reveals the moral hazard of a bailout. In that case, a representative banker takes advantage of the safety net to expand her investment. In this section, the same moral hazard continues to apply.

Moreover, Proposition 2.7.1 shows that bankers have an added incentive to coordinate their investments. Without bailouts, bankers may want to distance themselves from one another so that they can avoid a systemic crisis and its associated liquidation costs. With a social safety net, bankers intentionally concentrate their investments to enjoy higher upside profits, leaving taxpayers to pay more per bank on the downside.

The analysis in this section further emphasizes the need for ex-ante regulation. In our particular setting, the key welfare-reducing factors are the costs of liquidation and bailout. Without the government’s ex-post intervention, bank liquidation both inhibits the bankers’ economies of scale and imposes a dead-weight loss on social welfare. While ex-post bailouts prevent bankruptcy and enable bankers to invest more, they come with their own cost as well.

To eliminate both of these costs, in Section 2.6 we show that it is optimal to employ a policy combination of size regulation and CoCos. The result holds in this section, as the same requirement forces banks to raise enough CoCos to
avoid liquidation and bailouts. In other words, our policy proposal in Proposition 2.6.1 is robust to our extension of two banks and correlated risk. In this case, bankers will concentrate their investments to enjoy the added benefit, $\Delta$, but at no cost to taxpayers. Social welfare also reaches the first-best level.

2.8 Conclusion

In this paper, we provide a simple model to analyze the TBTF problem and related policy responses. In our view, TBTF emerges naturally from the time inconsistency of bailout commitment. By taking an excessive investment, bankers increase the chance and benefit of a bailout in a downturn, which creates the TBTF issue. To discourage banks from becoming too big, the government can employ a range of policies and policy combinations, including a capital requirement, cap on size, tax on size, and an additional regulatory capital (i.e., CoCos).

Our paper has shown the advantages and disadvantages of each individual policy. For example, a capital requirement is ineffective in preventing TBTF because our representative banker has limited equity capital. Size regulation, which directly restricts excessive investment, limits the banker’s efficiency in the economy. Therefore, policy responses to TBTF should not be only about size but also about the banks’ economies of scale. As we have shown, while size regulation is necessary, allowing CoCos to serve as an additional regulatory capital can both restrict TBTF and restore efficiency. Such a policy combination induces effective risk-sharing between bankers and investors and helps to achieve a higher level of social welfare.
Chapter 3

A Historical Walk Through Financial Reforms: Lessons for the Future

This chapter is joint work with Dang Du at the Federal Reserve Bank of St. Louis. The views expressed herein are solely those of the authors and do not necessarily reflect those of the Federal Reserve Bank of St. Louis.

3.1 Introduction

Historically, financial reform in the United States has been a piecemeal, fragmented process that reflects evolution in legislation, a patchwork of rules and regulations accompanied by social, political, and economic transformations. Rules and regulations are inherently reactive, because they reflect Congressional intent, and Congress is by nature a reactive institution. Based on major reforms that have left (or are expected to have) lasting impacts on the financial landscape, we have observed that incremental reform and intentional fragmentation have significantly increased complexity in the system. In particular, each
time a new and extensive set of rules and regulations is grafted on to the existing apparatus, it creates room for regulatory and supervisory leakages to evolve alongside financial innovations.

After the most recent crisis (2007–09), for instance, Congress approved the Dodd–Frank Wall Street Reform and Consumer Protection Act (Dodd–Frank Act) of 2010, which is one of the most ambitious financial overhaul packages since the Great Depression. It reflects the intricacy of the regulatory system, the score of compromises between committees and party leaders, and the delicate balance between regulatory tightening and supporting tepid economic recovery. By some measures, the aggregate word count in the Dodd–Frank Act, which refers to financial restrictions, dwarfs any legislation in modern history. In addition, it placed a huge burden on a web of both old and new regulatory agencies and bureaus to issue a slew of new regulations, reports, and studies, many of which remain incomplete to this day.

In a larger context, although this comprehensive reform brought an end to more than three decades of financial deregulation, the Dodd–Frank Act is not the first legislation that has attempted to correct failures in regulatory oversight, transparency, and credit excesses in the financial system. Some specific issues that the Act addresses involve the lack of stability and resilience of short-term funding markets, ineffectiveness of the federal regulatory regime, and existence of unregulated bank-like financial intermediaries. These issues are not new. They merely manifest in different ways, gradually transforming, adapting, and transcending in order to take advantage of the contemporary regulatory environment. For instance, some of the main factors that caused the Panic of 1907
included the frequent and widespread runs on deposits, lax and uneven regulatory standards at state and federal levels, and rapid growth of unregulated trust companies. Accordingly, they all played a major role in policy discussions and debates that eventually led to the passage of the Federal Reserve Act in 1913.

To shed light on the reciprocal and recurring relationship between regulatory overhauls and crises, this paper provides a review of five crucial financial reform initiatives, while placing special focus on regulatory leakages. The five reforms include the following: (1) the National Banking Acts of 1864; (2) the Federal Reserve Act of 1913; (3) the Banking Act of 1935; (4) the deregulation of the 1980s; and finally, (5) the Dodd–Frank Act of 2010. This paper focuses on these reforms because not only did Congress pass each of them specifically in response to a major financial crisis, but each reform also has significantly shaped the financial industry in the United States and will continue to do so for decades to come. While the proximate causes of each crisis are debatable, the connection between each reform and the subsequent crisis is due to a common factor that has existed in the United States since the Civil War. By examining these reforms and placing them within a larger evolutionary context of the financial regulatory development, this paper traces the developments of regulatory leakages that have led to the rise—in various forms—of unregulated short-term funding markets, ineffective regulatory regimes, and shadow banking.

First, as regulations change over time, banks continually search for new and cheaper sources of liquidity. Before the Civil War, banks issued their own currency in order to finance their loans and investments. With some 1,600 banks, nearly 10,000 types of currency notes were floating in the system in 1863 with wild fluctuations in both quality and quantity, thereby causing considerable
frustration to the general populace, financiers, and politicians alike. After the creation of a uniform currency in 1864, banks turned to demand deposits and interbank balances for short-term funding, the fragility of which played a major role in many financial crises over the next 65 years. After the Great Depression and two subsequent decades of conforming to the Banking Act of 1935 regarding the new and stringent regulation of deposit issuance, the financial system began to see an incredible growth in wholesale funding through money market mutual funds (MMMFs) and other financial intermediaries. This trend has continued until the present day. According to the Financial Crisis Inquiry Commission (2011), banks’ overreliance on wholesale funding became one of the main reasons underlying the most recent financial crisis (2007–09). From this perspective, regulations targeting short-term funding are likely to remain a key factor requiring further attention from lawmakers.

Second, regulatory failures have been a result of regulatory leakages that trace back to the origin of U.S. banking. In 1864, when Congress authorized the Office of Comptroller of the Currency (OCC) with the regulation of the newly established national banking system, it left existing state banks under the supervision of state authorities. Between 1865 and 1913, the OCC and state regulators repeatedly lowered regulatory standards, including capital requirements, barriers to entry, and reserve requirements, in an effort to attract and retain bank membership. Between 1913 and 1930, in the presence of the Federal Reserve System (Fed), whose members consisted of both state and national banks, state and federal regulators once again competed to extend branching rights and reduce portfolio restrictions for banks in their jurisdiction, thereby allowing banks to become larger and riskier. The regulatory changes of the 1930s
established a new host of rules that significantly reduced financial firms’ ability to take risks; however, at the same time, they allocated considerable discretion in the hands of regulators. While the new regulatory regime prevented the financial system from adapting to the changing macroeconomic environment in the early 1970s, regulators exercised a great deal of forbearance that contributed to the subsequent Savings and Loan Crisis of the 1980s. Thereafter, a trend of deregulation ensued that eventually led to the 2007–09 crisis.

Finally, during each period that led up to a major financial maelstrom, the system would witness the rise of new types of unregulated (or lightly regulated) financial intermediaries that the existing regulatory regime could not address and that lawmakers and policymakers could not bring under supervision in time. In the most recent crisis, these included hedge funds, structured investment vehicles (SIVs), special purpose entities (SPEs), money market funds, and other non-bank financial institutions, which collectively make up the current “shadow banking” system. Although the term “shadow banking” did not become popular until Paul Allen McCulley, an American economist, coined it in 2007, the concept can be retroactively applied to various financial entities in the past. Examples include “wildcat” banks before 1864 that issued highly-leveraged currency notes in remote areas, bank-like trust companies in the early 1900s that had less restricted investment scopes than banks, and investment affiliates that commercial banks created to finance stock speculations before the Great Depression.

While the financial system has changed considerably over time and a number of major reforms have continued to reshape the regulatory apparatus since the Civil War, the policy debates regarding the status and future of financial
regulation are current and pressing. Although the Dodd–Frank Act of 2010, in many ways, has provided a broad and compelling foundation for policymakers to address existing regulatory leakages, a number of previous legislations have arguably attempted to accomplish similar goals. The 2007–09 crisis and the very passage of the Dodd–Frank Act imply that previous reforms must have failed in certain aspects. Thus, examining the development of the financial regulations through a historical lens is imperative to trace and understand any deep-rooted regulatory failure that may still exist in the system today.

The remainder of this paper is as follows. Section 2 discusses the historical context of the National Banking Act of 1864 and the resulting weaknesses of the National Banking Era. Section 3 highlights the leading causes of the Panic of 1907 with respect to regulatory leakages and discusses how the Federal Reserve Act addressed (or failed to address) them. Section 4 discusses the new challenges that arose during the early years of the Fed, and the subsequent reforms that the Banking Act of 1935 imposed on the system. Section 5 discusses the deregulation trend beginning in the 1970s. Finally, Section 6 discusses the Dodd–Frank Act of 2010.

### 3.2 The National Banking Era, 1863–1913

#### 3.2.1 A Historical Context

For nearly nine decades after the Declaration of Independence was ratified, the United States was without a national banking system and a uniform currency. In place was a collection of state-charted banks that issued their own money under the supervision of individual states. As regulatory standards differed
from state to state, the quality of these banks and their currency notes varied greatly throughout the nation. This caused much inconvenience to the federal government and the general populace as well as hindered the growth of interstate commerce. By the beginning of the Civil War, a large number of southern representatives who typically fought to preserve states’ rights withdrew from Congress, whereas the federal government scrambled to find an effective way to finance the war. Congress took the opportunity and passed the National Currency Act of 1863, which was the first major financial legislation in the United States.

One important motivation for the law lies with New York’s 1838 Free Banking Act, the passage of which reflected the mentality of the lawmakers at the time. Prior to this Act, bankers were required to be incorporated by the state legislature. In 1818, New York passed a law that prohibited, with few explicit exceptions, individuals and corporations from banking activities, including taking deposits and issuing currency notes. The reasoning at the time, as Hammond (1936) documented and explained later, was that banks were naturally monopolistic because they handled the state’ finances and controlled the circulating notes. Thus, it was necessary to limit the incorporation of new banks to the fullest extent. The fallacy soon became apparent as the 1818 Act not only failed to inhibit monopolies in the banking sector but also led to an “extended system of corruption” (Hammond, 1848). A measure that established a new type of legal entity—free banking associations—and abolished virtually all restrictions in the banking sector was then passed with two-thirds majority on April 18, 1838.
New York’s 1838 law marked the beginning of a new age that modern historians now refer to as the “Free Banking Era.” For New York in particular, the law allowed the state’s banking system to unleash its true potential. According to Bodenhorn (2008), although New York was perceived as the commercial center of the nation long before 1838, its financial sector was “inefficiently small” compared with that of Massachusetts and Pennsylvania. After the passage of the free banking law, however, the number of banks in New York almost doubled within three years, quickly surpassed the numbers in the other two neighboring states, and became significantly larger over the next two decades (Bodenhorn, 2008). In addition, New York required banks that issued currency notes to deposit state bonds with the state authority and redeem notes at par with specie, which protected the holders from losses that may arise in bank failures. Coupled with strict regulatory enforcements and sound supervisory practices, New York’s banking sector during the subsequent period was markedly stable and robust (Rockoff, 1974; Rolnick and Weber, 1983; Hammond, 1957).

However, while historians generally concur on the success of the free banking law in New York, the overall legacy of the Free Banking Era has been the subject of numerous contentions. During this era, 17 other states passed some form of free banking laws that largely mirrored the one in New York. The popularity of the free banking laws contributed significantly to the substantial growth of the banking sector during this period (Sylla, 1970; Klebaner, 1990). Meanwhile, others have expressed their resentment of these laws, because they enabled “wildcat banks” to roam freely, thereby causing instability to the system and significant losses to the people (Knox, 1903; Rockoff, 1974; Luckett, 1980). Rolnick and Weber (1982, 1983) provide a mitigating viewpoint, arguing that
although bank failures occurred with a high frequency during this era, banks performed no worse than firms in other sectors and that losses to depositors and noteholders were not entirely due to fraud but were partly due to the falling prices of state bonds. Furthermore, Ng (1988) provides a state-by-state examination of bank entries among the states with free banking laws and concludes that, except in New York, “free banking laws did not generally lower barriers to entry, increase competition in the banking industry, or increase social surplus” (p. 886).

The Free Banking Era, in any case, left one clear footprint on the pre-Civil War financial landscape in the United States: a significant number of banks, each with a number of currency notes in different denominations. By 1861, there were 1,601 banks with more than 10,000 types of currency notes across the nation (Davis, 1910; White, 1983). This was a source of great frustration to the general populace, financiers, and politicians alike. In his 1910 report to the National Monetary Commission titled *The Origin of the National Banking System*, Davis documented a series of newspaper articles, speeches by state governors, and reports by senators and congressmen to emphasize the problematic nature of existing banknotes. These include the sheer number of different brands, wild fluctuations in both quality and quantity, and accompanying counterfeits. In his 1907 book, *Jay Cooke: Financier of the Civil War*, Oberholtzer documented the opinion of Jay Cooke—one of the most prominent financiers of the time—who also expressed similar concerns regarding the banknotes.

Amidst the chaotic scene arising from the 10,000 different types of banknotes, the need for a uniform currency became more apparent and the demand stronger over time. In 1849, Millard Fillmore, the Comptroller of the State of New York,
suggested that a uniform currency issued on the basis of U.S. bonds would be a “great convenience” to both the government and the people (Davis, 1910, p. 10). In his 1856 inaugural address as the Governor of Ohio, Salmon Portland Chase noted that “a sound and sufficient currency [was] indispensable to the welfare of every civilized community.” In 1861, Mr. Chase, now as the Secretary of the Treasury, wrote his annual financial report to Congress, urging the adoption of a uniform currency. He argued that the people “in their ordinary business would find the advantages of uniformity in currency; of uniformity in security; of effectual safeguard, if effectual safeguard is possible, against depreciation; and of protection from losses on discounts and exchanges” (p. 19).

The National Currency Act, henceforth, came under serious consideration by Congress in early 1863. Notably, in July 1862, Congressman Samuel Hooper introduced a bill that already contained similar language and provisions as the final Act. However, the bill was eventually defeated in August. For the ultimate purpose of a single uniform currency, his bill sought to eliminate all state banknotes via a two-percent tax. Although it is unclear whether this particular provision remained in the bill when Congressman Hooper reintroduced it in the House in January 1863, the version that Senator John Sherman (after whom the final Act was later named) introduced in the Senate in early February 1863 did not mention state banknotes at all. According to Davis (1910, p. 84), the reason was that the Senate was empathetic to state bank stockholders who supported the Union government during the onset of the Civil War. By Senator Sherman’s persuasion and Secretary Chase’s “ultimatum,” the Senate passed the bill in less than 10 days of its introduction (Davis, 1910, p. 88). The House followed suit, passing the bill on February 20th, which President
Abraham Lincoln signed into law on February 25, 1863.

The hastened passage of the bills through Congress was not, however, due to the need for a uniform currency. To persuade Congress, Secretary Chase, Senator Sherman, Congressman Hooper, and others relied on the urgency of war finances. By passing the proposed bills and creating currency notes backed by U.S. bonds, they argued, the government would be able to raise additional debts to meet the rising costs of the on-going Civil War (Davis, 1910; White, 1983; Hammond, 1957; Champ, 2007). According to a report by the U.S. Congress (1880), the government’s ordinary revenue was merely $52 million in 1862, whereas the expenditure amounted to just below $570 million—of which $400 million was war costs (U.S. Congress, 1880). In his 1862 annual finance report to Congress, Secretary Chase pleaded that if a currency backed by U.S. bonds was created, the demand of U.S. bonds would increase by at least $250 million in no time. Later, when the bill was in the Senate, he went further and claimed that without its passage, he would be unable to raise additional debts (Davis, 1910, p. 83). Similarly, Senator Sherman, when being asked about the benefits of the bill, stated that “The first benefit is, there is a market furnished for the bonds of the United States” (Davis, 1910, p. 109).

Given the situation at the time, neither the House nor the Senate had any choice but to pass the bills. In 1862 alone, Congress had already passed the Morrill Tariff Act, the Legal Tender Act, and the Internal Revenue Act, all in an effort to raise the government’s revenue through customs, taxes, and even money printing (“greenbacks”). By the end of 1862, Congress must have realized that the war would cost more than the government’s existing finances could handle. In fact, while the government’s revenues stagnated around $400
million a year, its expenditures rose rapidly to $2 billions dollars a year by 1865, with the annual war costs rising to more than $1 billion (U.S. Congress, 1880). Had Congress not passed the National Currency Act in early 1863, the outcome of the Civil War—and the history of the United States—may have been very different.

3.2.2 National Banking and Its Weaknesses

The Sherman Act of 1863, as described above, passed through Congress and became law in a matter of days. It was no surprise then that, with such a hastened passage, the Act was “imperfect and incomplete” (Office of the Comptroller of the Currency, 2011, p. 5). A little more than a year later, on June 3, 1864, Congress repealed and replaced the Act with another act of the same name. The latter act—the National Banking Act of 1864—consisted of largely the same language and contents as the original one, save for a few changes. For the next 50 years, as the national banking system took form and evolved, regulatory leakages also became apparent.

The first form of regulatory leakage occurred through bankers’ increasing dependence on interbank balances that made the financial system vulnerable to systemic liquidity crises. While the National Banking Act of 1864 succeeded in its primary objective of creating a uniform currency, it also made the currency highly inelastic by placing a cap on its issuance and tying its supply to the availability of U.S. bonds. In addition, the Act created a pyramiding reserve structure that consisted of three types of banks (country, reserve-city, and central-reserve-city) with pyramidal reserve requirements. As the inelasticity of the currency caused all level of banks to shift their short-term funding sources
from currency to demand deposit issuance, the reserve structure encouraged the concentration of reserves in large cities, thereby creating a net of interbank balances that large city banks depended on for short-term financing and small country banks depended on in times of crisis. Consequently, a small liquidity shock to one region, then, could quickly erupt into a nationwide systemic crisis.

The inelasticity of the currency was inherent in its design and could not be fixed easily. As a result of the 1864 Act, not only did national banks have to deposit U.S. bonds at the Treasury to guarantee all of their notes, but they also could issue only 90 percent of the par or market value (whichever was smaller) of the deposited bonds. Hence, the supply of U.S. bonds, which depended solely on the financing needs of the government, effectively served as a constraint on the supply of currency notes. Despite the abundance of U.S. bonds during the Civil War, the emergence of government budget surpluses beginning in the 1880s seriously threatened the money supply. As early as 1882, Comptroller John Jay Knox noted in his annual report that if the government continued to pay off its debt as it had the three years prior, “all of the interest-bearing bonds [would] soon be surrendered and canceled” (p. 19). In addition, after a series of financial and economic crises throughout the latter half of the 19th century and at the beginning of the 20th, the National Monetary Commission, which was created by the Aldrich-Vreeland Act of 1907 to study the nation’s banking system, concluded the following in 1912:

“Of our various forms of currency the bank-note issue is the only one which we might expect to respond to the changing needs of business by automatic expansion and contraction, but this issue is
deprived of all such qualities by the fact that its volume is largely dependent upon the amount and price of United States bonds.” (p. 7)

Furthermore, the reserve structure was pyramidal and highly immobile. The 1864 Act, in addition to imposing a minimum reserve requirement on all national banks, identified three types of banks, which included country banks, reserve-city banks, and central-reserve-city banks. While the requirement for the first type was 15 percent, it was 25 percent for the other two. Moreover, a country bank could keep 50 percent of its reserve at a reserve-city or a central-reserve-city bank, while a reserve-city bank could keep 50 percent of its reserve at a central-reserve-city bank. Although the requirement was not small at each level, the total amount of reserve could be deceptively small in comparison with the total deposits because of the pyramidal characteristic. In addition to the three-tiered reserve system, three types of vault cash could be used to meet liquidity demand. The first resided with country banks, which were scattered around the country. The second resided with banks in 47 reserve cities, whereas the third with banks in three central reserve cities: New York, Chicago, and St. Louis. Although the total amount of cash reserves was significant, Johnson (1977) noted that the “economic value of this reserve was largely mitigated because it was so spread out” (p. 14).

Under these new designs, interbank deposits became necessary as well as profitable in the short run for bankers. When local economic conditions fluctuated, banks’ finances oscillated from having idle funds to facing high liquidity demand. Previously, because banks could issue or withdraw currency notes

\footnote{Some $1.5 billion in 1914, by authors’ calculation based on Watkins (1929)}
with little restriction, currency issuance served as the main source of short-
term funding for the financial industry and the economy at large. This could
no longer be achieved after 1864. Thus, bankers maintained correspondent ac-
counts with other banks so that they could earn interest on their surplus funds
and borrow when in need (Watkins, 1929). Interbank balances, which were
largely nonexistent in 1864, grew rapidly over the next 60 years. According to
Watkins (1929), national banks held (in total) $1.243 billion of deposits that
belonged to other banks in 1902—a figure that rose to $2.168 billion just 12
years later. Of the total interbank balances, country banks’ deposits accounted
for roughly 46 percent, while reserve-city banks for another 40 percent. More-
over, bankers’ balances concentrated heavily in only three cities: New York,
Chicago, and St. Louis. In 1914, for instance, the three cities together held
approximately 40 percent of all bankers’ deposits, whereas New York alone held
more than one-quarter.

However, the system of interbank balances was highly fragile and prone to
widespread runs. If a liquidity crisis occurred in a given area, local banks had
to meet local liquidity demand by drawing their balances from correspondent
banks, which, in turn, had to call on their balances at other correspondent banks
and so forth. The chain of events almost always ended at the central reserve city
banks, particularly those in New York. As New York central-reserve-city banks
were forced to liquidate their assets and recall their loans, the initial shock to a
small rural area, which could be easily localized in modern times, may have had
a ripple effect throughout the banking industry, stock market, and real economy
as a whole. In the absence of deposit insurance and a lender of last resort, the
National Banking Act of 1864 that fixed the wild fluctuations in the value of the
currency opened up the financial system to interbank lending, which became the source for previously unobserved systemic liquidity crises. During the short period between 1864 and 1913, a total of five major financial crises occurred, all of which featured widespread bank runs, causing severe disruptions to the financial system and the national economy (Sprague, 1910).

The second form of regulatory leakage manifested through the parallel existence of state and federal regulators and the competition between them. While the 1864 Act established the national banking system, it made no mention of state banks, except for a provision allowing state banks to convert. As a result, national banks were under the supervision of the newly-created OCC, while states authorities chartered and regulated state banks. With the 1864 Act’s failure to address state banks, the history of financial development in the United States was marked with the inception of the duality of banking that has lasted until the present day. According to Davis (1910, p. 100-102), the framers of the National Banking Act expected state banks to convert to national charters. Reality fell short of their expectations as the number of state banks, which was at 1,492 in 1862, stood at 1,089 in 1864 (White, 1983). State banks continued to issue their own notes, while most new national banks were small in comparison with their state counterparts (Davis, 1910). National banking supporters saw a brief triumph when the number of state banks fell drastically to 297 within a year of the 1865 Amendment (which placed a 10 percent tax on state currency notes) and remained substantially below the number of national banks for the next 25 years. Nonetheless, by the end of the period in 1913, the country had more than 14,000 state banks—twice the number of national banks at the time (White, 1983).
The resurgence of state banking was due almost entirely to the competition in laxity between state and federal regulators. Accepting the fact that their banks could no longer issue notes profitably, state legislators sought alternative channels to prop up state banks’ viability and competitiveness. Beginning as early as 1870, states shifted their focus to lifting regulatory burdens for their banks, including allowing intra-state branching, lowering capital and reserve requirements, reducing portfolio restrictions, and in some cases, neglecting the supervision and examination of banks. In his 1983 book, *The Regulation and Reform of the American Banking System, 1900–1929*, White detailed the changes in different states while noting that the OCC—the sole federal regulator at the time—responded to the resurgence of state banking via the same channels, creating a “race to the bottom” in regulatory standards throughout the National Banking Era. By 1913, state banks in most states faced lower capital and reserve requirements than national banks; could invest in riskier and longer-term assets; could geographically diversify; and were generally subject to laxer regulatory standards and less supervision and examination. In addition, national banks in 1913 enjoyed across-the-board lower regulatory burdens than those in 1865.

Finally, the third regulatory leakage of the National Banking Era occurred through the rise of unregulated shadow banking. Because the National Banking Act of 1864 neither prevented nor regulated the existence of other types of depository institutions, a gateway for shadow banks to propagate in the system was open. Consequently, bank-like trust companies emerged. Trust companies, which were originally designed to serve the wealthiest members of society, were similar to typical commercial banks in many ways. Particularly, they could
receive deposits in trust and engage in a wide range of investment activities traditionally handled by banks. Gradually, they began to compete directly with commercial banks for deposits and expand their business into lending on collateral securities and buying commercial papers, just like major city banks did at the time. Because they were state chartered, trust companies enjoyed the same regulatory advantages as state banks. In particular, they were generally not subject to reserve requirements; could invest in a wider range of assets, including real estate and stocks; and could have virtually unrestricted portfolio allocations.

Because of the uneven regulatory standards applying to banks and trust companies, shadow institutions grew rapidly in both number and sizes. Between 1886 and 1900, trust companies increased sixfold in number and in assets (White, 1983, p. 39). In the seven years leading up to 1907, the number of trust companies grew from 290 to 794, while their assets grew from $1.3 to $3 billions. In New York alone, trust companies held $1.3 billion in assets in 1907, in comparison with $1.8 billion held by national banks (Moen and Tallman, 1992). Thus, unlike previous crises in which runs on deposits began with and more severely affected national banks, the Panic of 1907 originated from trust companies in one corner of New York and quickly spread throughout the entire banking system and the country.

Because of these major regulatory leakages in the banking system—the rise of interbank balances, regulatory competition, and shadow banking—the National Banking Era experienced more frequent systemic liquidity panics and recession than any other period in history. Within 50 years (1864–1914), there were five major financial and economic crises that both originated from and were
amplified by these factors (Sprague, 1910). In response, the nation entered a
new string of debates regarding financial regulatory reforms, the result of which
was the eventual passage of the Federal Reserve Act of 1913.

3.3 The Fed: History, Politics, and Leakages

3.3.1 A Prelude to the Federal Reserve Act

The first decade of the 20th century was a remarkable period that brought an
end to the National Banking Era. In 1900, Congress passed the Gold Standard
Act, which put to rest national debates on the value of the U.S. currency and
the use of specie payments. Prior to the Act, the aforementioned weaknesses of
the national banking system were not apparent, because each crisis was often
associated with the uncertainty of the gold standard that led to speculation and
unexpected flows of gold in and out of the country. Although these weaknesses
culminated in the Panic of 1907, the circumstances of this panic, as well as
subsequent events in the following five years, focused national debates regarding
financial reforms on one particular issue: combating the system’s vulnerability
to systemic liquidity shocks via the establishment of a central reserve institution.

The Panic of 1907, like previous crises, was a classic liquidity crisis that
the rigid national banking system at the time could not weather effectively.
Beginning with a failed attempt by the “copper king,” F.A. Heinze, to corner
the stock of United Copper, investors began to run on the largest trust company
at the time, the Knickerbocker Trust. While the company was solvent by all
available evidence, the run started on the basis of a rumor that Knickerbocker’s
president had a close connection with Heinze (Tallman, 2012). As the run forced
the company to suspend payments, many other trust companies, and eventually commercial banks in New York City, experienced widespread withdrawals of deposits. Despite efforts to provide the market with the necessary liquidity by the New York Clearing House, the U.S. government, and J.P. Morgan (the most influential financier at the time), panics continued to spread throughout the city. By October 26, 1907, after 10 days of panics and runs, commercial banks in New York began systematic restriction of cash payments, which commercial banks in the rest of the country soon emulated. [For detailed discussions of the Panic, see Sprague (1910); Moen and Tallman (1990); Bruner and Carr (2008); Strouse (1999); Wicker (2000)].

Notably, the actions of the New York Clearing House during the Panic of 1907, which were aimed at subduing and preventing the widespread nature of the crisis, gained a great deal of public attention. Beginning in the 1850s, the United States began to see a new form of central banking that arose to fill the need of coordination among national banks. In 1865, 52 banks in New York City, each of which previously had to deal with one another via costly methods of maintaining high balances and a system of messengers, formed the NYCH—the first clearinghouse in the nation. The NYCH not only helped clear interbank claims but also imposed regulation and supervision with higher standards than the federal requirements (White, 1983, p. 74–75). By the turn of the century, the NYCH was functioning as a de facto central bank for New York City. As such, during the Panic of 1907, NYCH served as a lender of last resort when extending credit to many member banks, issuing clearinghouse certificates to increase liquidity in the market, and imposing restrictions on cash payments throughout the city to prevent liquidity drains.
However, the NYCH was not without its limitations. First, it was, by nature, a collection of member banks; therefore, it had little interest in protecting the stability of the financial market at large. While it provided emergency liquidity to member banks and mitigated the spread of panic during previous crises, during the Panic of 1907, the NYCH repeatedly refused to assist the Knickerbocker Trust Company, a non-member (Tallman, 2012). The NYCH’s action, albeit justified as safeguarding the reserves for its members, failed to quell the seed of the panic, which subsequently led to a liquidity shortage for its own members. Second, the NYCH, which was the largest clearinghouse in the country, found itself short of the necessary reserves to tackle the panic once it was set in motion. After 10 days of employing all available means, including raising additional capital from private financiers and receiving assistance from the government, the NYCH resorted to restrictions of cash payments among all of its members, which was highly disruptive to the local economy (Sprague, 1910, p. 273). Finally, the NYCH’s actions were confined to the banking sector in New York City and did not address the concerns of banks in other locations. The NYCH’s suspension of payments, while protecting banks in New York, spread the panic to other banks around the country, which not only suspended payments in their own banks but also began to draw down their interbank balances. This, amplified by the pyramiding reserve structure, significantly reduced the cash reserve in the entire nation. Regarding the nation’s reserves, White (1983) recorded the following:

“Between August 22 and December 3, 1907, country banks increased their cash by 23.4 percent and reduced their correspondent
balances by 13.5 percent while cutting loans by 3.5 percent. This strained the reserve-city banks, which lost 14.6 percent of their cash and responded by drawing their interbank deposits down by 17.3 percent and reducing loans by 4.7 percent. In the center of this maelstrom, New York City banks lost 19.1 percent of their cash, as their correspondents drew down on their balances...” (p. 70)

Realizing the need for financial reform, Congress quickly passed the Aldrich–Vreeland Act, the single most important and lasting impact of which was the creation of the National Monetary Commission. Through the Act, Congress charged the Commission, which was composed of nine members from each congressional house, with the task of studying the national banking system, identifying its weaknesses, and recommending changes and reforms for the system. By 1912, the Commission produced a total of 30 reports on an expansive range of topics, including national and state banking laws, financial laws, and the international banking and currency systems of other industrialized nations. These reports emphasized heavily on the need for a centralized institution that could (1) facilitate the elasticity of the currency via lending as the last resort; (2) concentrate and mobilize reserves; and (3) provide uniform standards in regulations and supervision across all commercial banks (National Monetary Commission, 1912). In other words, the Commission argued for the necessity of a central bank.

Based on the Commission’s recommendation, Congress considered the earliest plan, the Aldrich plan, in 1911; however, the plan did not go very far. The Aldrich plan called for a single central bank, a “National Reserve Association,”
with branches throughout the country and the power to issue currency and re-
discount commercial papers of member banks. It would be under the control of a board of directors, most of whom would be bankers. Unfortunately, the plan received little support from the public and came under heavy attacks from Democrats, who had just won Congress. By 1912, the Pujo hearings before the House Banking and Currency Committee convinced the American public that the country’s finances were under the control of a small group of people on Wall Street, the “money trust” (Bruner and Carr, 2008; Bordo and Wheelock, 2013).

In addition, Democrat Woodrow Wilson, who had just won the 1912 presidential campaign, vowed (ironically) to reform the financial system without the creation of a central bank. By the end of 1912, the Aldrich plan was effectively dead.

Upon assuming the presidency, President Wilson relied on Representative Carter Glass and an expert adviser, H. Parker Willis, to draw up a new plan in 1912 that focused on resolving the problems with immobile reserves and inelastic currency. The Glass–Willis plan proposed a system of 20 or more privately controlled reserve banks that would issue currency against commercial assets and gold as well as perform central banking functions. To increase government oversight, President Wilson added a central board that would control and coordinate the work of the regional reserve banks. According to Johnson (1977), while this plan formed the basis for the Federal Reserve Act, it continued to face opposition from both sides. On one hand, bankers and conservatives believed that the plan gave the government too much power. On the other hand, the “agrarians” from the West and South thought that the plan left too much power in the hands of private bankers and gave the government too little oversight.
At this point, the focus of policy discussions turned into debates regarding the control of the new system. In the spring of 1913, President Wilson sought advice from Louis D. Brandeis, whose adjustments would then give the Glass–Willis plan a fighting chance in Congress. First, to please the agrarians, the government would retain exclusive control of the central board. Second, to please the bankers, the central board would have less authority over the discount rate, leaving regional banks with more power and independence. In addition, to facilitate the relationship between the board and regional banks, bank representatives would form a federal advisory council to consult the board on relevant policy issues. With these changes, Representative Glass and Senator Owen introduced the Federal Reserve bill in Congress in June 1913.

Opposition persisted for the next six months as both houses continued to debate and modify the bill. While the House passed a version of the bill in mid-August, one of the most significant last-minute challenges to the bill came in October. Answering an invitation from the Senate, former Assistant Secretary of the Treasury Frank Vanderlip, who was then the President of the National City Bank of New York, unexpectedly introduced a new proposal in the Senate that called for a single Federal Reserve Bank with 12 branches, entirely under the government’s control. As President Wilson stepped in to voice his strong opposition, some elements of the Vanderlip plan were incorporated into a modified Federal Reserve bill, which the Senate passed on December 19.

On December 23, President Wilson signed into law the Federal Reserve Act of 1913, thereby establishing the Federal Reserve System (Fed) in the United States. He noted that the bill would “be of lasting benefit to the business of the country” (Johnson, 1977, p. 34). By February 1914, 99 percent of national
banks joined the system to retain their charters, despite earlier warnings that they would not (Willis, 1923). By April 2, 1914, the Reserve Bank Organization Committee, after much debate to wind down the power of New York, announced the district lines and 12 reserve cities. On August 10, 1914, after months of struggling to organize the Federal Reserve Banks (FRBs) and Board, the Board was officially sworn in, with Charles S. Hamlin as the first Governor. By November 16, 1914, all the regional FRBs began conducting business.

3.3.2 The Fed’s Early Years, 1913–1933

The passage of the Federal Reserve Act of 1913 was a major turning point in the development of financial regulations in the United States. For the first time, the country had a large banking system with the presence of a central bank, much like all of the other developed nations of the time. However, the earlier design of the Fed was not without its flaws. As discussed in the previous section, the debates on reform focused heavily on the establishment of a central reserve system and turned eventually on who would hold control of it. Consequently, the financial system was still exposed to systemic shocks, shadow banking was never addressed explicitly, and regulatory competition between state and federal regulators worsened. For the next two decades, particularly through the roaring twenties and the 1929–33 banking panics, these weaknesses of the Fed surfaced once again.

In its early years, the Fed enjoyed great success in providing liquidity to the market and managing the money stock. When Congress passed the 1913 Act, it authorized the Fed to deal in short-term commercial and agricultural papers, which was in accordance with the “real-bill doctrine” of the time (Bordo
The new system sought to encourage the use of these papers, increase their liquidity, and provide stability to the short-term funding market, much like the required use of U.S. bonds to back currency notes. However, the availability of commercial papers was already in rapid decline, as they accounted for 69 percent of loans in 1880 but only 35 percent in 1914 (White, 1983). When the United States joined World War I in 1917, the Fed employed an exception embedded in the 1913 Act to help finance the war. This exception allowed the Fed to deal in U.S. bonds, which the Fed began to buy in large quantities and to discount for banks at favorable rates. As a result, the Fed’s holdings of U.S. bonds rose from 4.3 percent in April 1917 to 55.1 percent a year later. By May 1919, 95 percent of discounts at the Fed were backed by U.S. bonds (White, 1983). Meanwhile, the monetary base grew at a rate of 11.2 percent in 1915, 15.2 percent in 1916, 20.6 percent in 1917, and 16.2 percent in 1918 (Friedman and Schwartz, 1963, Table 33). These actions significantly enhanced the U.S. government’s ability to finance and win World War I.

While the Fed’s action in time of war faced little opposition, it resulted in post-war inflation and brought the real-bill doctrine to the forefront of subsequent policy debates. Throughout the 1920s, for instance, Fed officials interpreted banks continuous borrowing in recession as being for the purpose of speculation in the stock market, and thus issued annual reports that discouraged banks’ use of the discount window (Meltzer, 2003, p.163). The reasoning was that when real economic growth was slow, the Fed’s credit, which was meant to facilitate loans to industrial and agricultural sectors, should be reduced. Not only the Board and several Bank governors but also Congress members and
much of the financial press joined the effort in supporting this view (Meltzer, 2003, p. 246).

Furthermore, as Congress grafted the FRBs onto the existing national banking system as “super clearinghouses”, the problematic nature of pyramidal reserve was never resolved. After the passage of the 1913 Act, the hierarchy of interbank deposits remained unchanged. Because the Fed could only lend to member banks with high-quality commercial and agricultural papers, rural and small city banks had to borrow from major city banks using papers that would not qualify at the Fed. Non-member banks, the number of which stood at 20,000 in 1920, did not have access to the discount window; thus, they had to rely on the system of interbank balances. Furthermore, while the FRBs acted as clearinghouses and enjoyed some success, private clearinghouses retained their dominant position in the market. In 1915, private clearinghouses cleared approximately $163 million, while the Fed cleared $5 million. By 1929, the Fed cleared $367 million; however, private clearinghouses cleared $713 million (White, 1983, p. 108–144). The reason was that the Fed imposed a strict guideline, which many banks, members and non-members alike, avoided by going through private clearinghouses.

At the onset of the Great Depression, although the market looked to the Fed to provide liquidity, the Fed’s decentralized structure hindered its ability to act as a lender of last resort. The financial system at that time was still exposed to systemic liquidity shocks, because bankers continued to depend heavily on interbank balances as they did during the National Banking Era. Only the Fed, as the nation’s super clearinghouse with virtually unlimited financial firepower, could have halted the spread of banking panics. However, as Friedman
and Schwartz (1963) argue, the Fed’s fragmented leadership was a major failure during the Great Depression. While each regional FRB could set its own policy, including setting the discount rate for their respective region and conducting open market operations, the Board could only approve or disapprove FRBs’ policies but could not compel them to take action. Consequently, some FRBs were more aggressive in responding to financial disturbances than others (Chandler, 1971, p. 233). For instance, the FRBs of Atlanta and New York, by intervening swiftly in their respective local markets in 1929 and 1930, were able to manage banking panics in their region better than those in other locations, including Chicago and St. Louis (Richardson and Troost, 2006; Bordo and Wheelock, 2013). During the first few years of the Great Depression, the tension between the Board and the New York Fed also ran high. Following the stock market crash in 1929, the FRB of New York responded quickly with expansionary open market operations. Yet, the Board disapproved, following which it rejected repeatedly the FRB of New York’s subsequent requests to lower the discount rate (Bordo and Wheelock, 2013). As Friedman and Schwartz (1963) argue in their classic book, A Monetary History of the United States, the Fed’s tightening of monetary policies in the middle of the 1929–33 banking panic was effectively the main cause of the subsequent Great Depression.

With regard to shadow banking, the Federal Reserve Act of 1913 neutralized the role of trust companies; however, it could not prevent the rise of other types of unregulated entities. Realizing the damaging impact that trust companies exerted during the Panic of 1907, the framers of the 1913 Act included a provision that encouraged trust companies to join the system: Fed-member trust companies would be subject to stricter regulatory and supervisory standards.
than before; however they would gain access to the Fed’s discount window. In addition, the Act also authorized commercial banks to conduct a trust business. These two changes in the system brought one part of the trust sector out of the shadows but left the rest growing into a new form. Particularly, the 1920s saw investment trusts—which pooled funds from individuals to invest in stocks, much like modern-day mutual funds—grow nearly 20 times in number (Carosso, 1970). Furthermore, as the 1920s featured a booming market, commercial banks in the traditional form could not compete with investment banks in the provision of funds to businesses. Particularly, investment banks could float new issues of bonds and stocks, which was outside the legal scope of commercial banks. To increase their competitiveness, commercial banks formed securities affiliates, which gave them access to all investment banking and brokerage activities. According to Peach (1941), the number of investment affiliates grew more than 11 times between 1922–1931, allowing commercial banks to gain 45 percent of the bond-originations market by 1929. Notably, both investment trusts and investment affiliates were on the outside of the existing regulatory apparatus, making them the new shadow banks. These entities played a major role in fueling speculation in the stock market, which lasted throughout the 1920s and led eventually to the Crash of 1929 (White, 1990, 2009). Needless to say, although the Federal Reserve Act of 1913 patched up one unregulated leakage, another one appeared.

Finally, despite the phrase “to establish a more effective supervision of banking” being part of the title of the Federal Reserve Act, the 1913 Act not only failed to implement any fundamental reform of banking regulation but also created additional interagency conflicts of interest. Competition in laxity between
state and federal regulators worsened, as the duality of banking continued to exist. Although state banks were encouraged to join the system, which offered access to the Fed’s discount window, the Act (and subsequent amendments) did not place any restraint on banks that refused membership. This left room for banks to shop between state and federal regulators, which in turn induced regulators to modify their standards and cater to banks’ interests. For instance, in 1915, 15 states lowered their reserve requirements in an effort to retain their hold on state banks. As the Fed responded by lowering its own reserve requirement, many more states reduced their requirements further and began to allow reserves to be kept in interest-bearing deposits at other institutions—a feature that the Fed did not allow. By 1929, this competition resulted in reserve requirements in many states being below the level that banks would have held in the absence of regulation (White, 1983, p.149). While the number of state banks that signed up for membership increased to 1,177 in 1929, the number of state banks that did not remained significantly higher, hovering between 15,000 and 20,000 throughout the Fed’s early years (White, 1983, p. 132).

Furthermore, the Federal Reserve Act, which charged the Board and the FRBs with the supervision of member banks, created strong friction between the Fed and the OCC. For instance, a state-chartered bank that applied for membership in the Fed could circumvent the Board’s approval (or disapproval) by seeking a national charter from the OCC. This was exactly what happened, as Governor William Proctor Gould Harding (1925) noted in his 1925 memoir, The Formative Period of the Federal Reserve System. In addition, while the Board had the authority to demand reports and examine member banks, the OCC was initially responsible for conducting examinations. Then, the Fed faced strong
complaints from state member banks, to which the OCC applied the same, more stringent regulatory standards being applied to national banks. It also did not help the Fed that, in the fight to retain its perquisite regarding the supervision of banks, the OCC refused to share its reports of bank examination with the Fed (White, 2013, p. 35). As tension rose between the Fed and the OCC, a 1917 amendment to the Federal Reserve Act stripped the OCC of its power to examine state member banks and gave it to the Board. Subsequently, to appease and retain state member banks, the Board slowly abated supervisory standards, including requiring less detailed information, abandoning surprise call reports, and reducing the number of call reports (White, 2013).

The weaknesses in the design of the Fed, as noted above, became apparent gradually throughout the Fed’s early years (1913–1933) and culminated in the Great Depression. The rise of a new breed of shadow banks contributed greatly to rampant speculation throughout the 1920s, while competition in laxity eroded regulatory and supervisory standards at both the state and federal levels. Furthermore, although the Fed was supposed to act as the lender of last resort in time of need, its decentralized and fragmented structure significantly hindered it from carrying out its policies, to the point of amplifying the 1929–1933 crisis. After the banking panic ended in March 1933 due to President Roosevelt’s actions, it was clear that the nation required another banking reform to restructure and strengthen the financial system.
3.4 Financial Reforms in the 1930s

After the 1929–33 banking panic ended, lawmakers began to contemplate a new design for the entire financial system that could prevent—or at least withstand—future crises. This effort resulted in a series of new laws that make up the most comprehensive reform ever seen. These laws include the National Housing Act of 1934, which established the national savings and loan industry; the Federal Credit Union Act of 1934, which established a federal system of not-for-profit, cooperative credit unions; the Securities Exchange Act of 1934, which established the Securities and Exchange Commission (SEC) to regulate secondary trading of securities; the Commodities Exchange Act of 1936, which regulated all commodities and futures trading activities; and the Banking Act of 1935, which was the farthest-reaching reform to commercial banking and to the financial sector as a whole. Altogether, these Acts defined the regulatory environment for decades to come.

The most important legislation that addressed regulatory leakages discussed in the last section was the Banking Act of 1935. Like all previous reforms, the Act was the result of lengthy political debate and compromise. An earlier version of the Act, which Congress passed in June 1933, was a combination of the bills proposed by then-Senator Carter Glass and Representative Henry Steagall while the panic was in full swing. As a result, both Senator Glass and Representative Steagall focused primarily on changes that were directly apparent.

For instance, the Glass bill, which the Senate passed in January, aimed at
eliminating under-regulated state banks, unifying all banks under the supervision of the Fed, and completely separating investment and commercial banking. Meanwhile, Representative Steagall wanted to provide federal deposit insurance to all banks, including both state and national. After a political compromise, Congress passed the Banking Act of 1933 with two major changes to the system—the creation of the Federal Deposit Insurance Corporation (FDIC) and the separation of investment and commercial banking. By the time Congress reconsidered the Banking Act in 1935, restructuring the Fed became the forefront issue. In an effort to centralize the Fed and make it more effective, an early proposal greatly increased the president’s authority and allowed the Fed to purchase U.S. bonds of any maturity. According to Bradford (1935), the proposal met with a “flood of protest and criticism,” with detractors detesting both the possibility of an inflationary environment and the departure from the real bill doctrine of the original Federal Reserve Act. Consequently, the Senate Committee on Banking and Currency held extensive hearings and made a number of amendments to the proposal (Bradford, 1935). Finally, after its passage through Congress on August 19, the Banking Act of 1935 became law on August 23 with President Roosevelt’s signature.

The most successful aspect of the Banking Act of 1935 was the new regulatory design of commercial banking, which protected the financial system from systemic liquidity crises for the next seven decades. First, the Act reorganized the Fed’s leadership structure, reduced the FRBs’ independence, and concentrated power into the Board. These changes, which gave the Board more control over discount rates at all FRBs and more power to set uniform monetary policies, enhanced the Board’s lender of last resort authority. Second,
the Act implemented a permanent program of federal deposit insurance, which was based on the temporary program of the 1933 Act. The FDIC, which received capital contributions from the government and the FRBs, was to provide deposit insurance for all Fed-member banks and qualifying state non-member banks. The model for deposit insurance was not new, as various states in the West and the South had already experimented with an FDIC-like system since the Panic of 1907 (White, 1983). However, the state systems, which depended largely on depositor confidence, began to fail during the 1923–24 economic recession and disappeared completely at the peak of the 1929–33 banking panics (White, 1983, p. 220–222). The advantage of the FDIC, therefore, lay mainly with the credit of the federal government—and American taxpayers. Since the establishment of the FDIC, “no depositor has lost a single cent of insured funds as a result of a failure.”

Third, the 1935 Act, by amending Section 19 of the Federal Reserve Act, forbade interest payments on demand deposits and granted the Fed the power to set interest rates on saving and time deposits. The rate ceiling regulation was written into Section Q of the Code of Federal Regulation and became known as the infamous “Regulation Q.” This change removed banks’ ability to compete for deposits and thus limit their tendency to invest in risky assets. Although the Banking Act of 1935 did not address bankers’ reliance of interbank balances, which was a major issue (as discussed in the previous section), the changes in the Act were sufficient to mitigate the system’s exposure to systemic liquidity shocks.

In order to address the threat of shadow banking, Congress redesigned the financial system as a whole. First, following the Banking Act of 1933, the 1935

\[\text{FDIC Website. } \text{https://www.fdic.gov/about/learn/symbol/}. \text{ Retrieved 04/01/2015.}\]
Act imposed a strict separation between investment and commercial banking, thereby forbidding depository institutions from trading on their own account or underwriting stocks and bonds. Consequently, commercial banks could no longer own and operate investment affiliates. Second, the Securities Exchange Act of 1934 imposed a wide range of requirements on all publicly listed companies. The requirements included the registration of any securities listed on stock exchanges, disclosure, proxy solicitations, and margin and audit requirements. The purpose of the Act was to increase transparency in the secondary security market by forcing companies to provide sufficient information to investors.

Third, the Commodity Exchange Act of 1936, which replaced the Grain Futures Act of 1922, extended federal authority over the regulation of commodity futures prices, prohibited market and price manipulation, and prevented fraud that may arise in brokerage firms. Similar to the Securities Exchange Act, this Act aimed at increasing transparency in the commodity market and bringing previously unregulated entities out of the shadows.

While the reforms in the 1930s were successful in many aspects, they failed to address regulatory competition, which was a crucial regulatory leakage in the system highlighted in the previous section. The duality of banking continued to exist, while the FDIC became the newest addition to the list of federal regulators of commercial banking. Furthermore, it was relatively easy for banks to change their charters and regulators. For instance, Scott (1977) reports 1,828 switches between 1950 and 1977, while Rosen (2005) reports 2,298 switches between 1977 and 2003. When a bank decides to leave a regulator, the regulator’s budget could be reduced by as much as 30% (Scott, 1977; Rosen, 2003; Blair and Kushmeider, 2006). As a result, competition between state and federal
regulators intensified, particularly via the exercise of their discretionary power (Scott, 1977). In addition, in an effort to promote homeownership in the United States, Congress passed the Federal Home Loan Bank Act of 1932 and the National Housing Act of 1934, which grafted the Federal Home Loan Bank System and the Federal Savings and Loan Insurance Corporation (FSLIC) onto the financial industry. This created a regulatory structure in the thrift sector that paralleled to the commercial banking sector. Consequently, regulatory competition between state and federal regulators in the thrift industry began, and this subsequently contributed to the Savings and Loan Crisis in the 1980s (Kane, 1989).

During the four decades after the Great Depression, because of the aforementioned reforms, the U.S. financial system was remarkably stable. On the liability side, banks did not have to compete for deposits because of both deposit insurance and interest rate ceilings. On the assets side, banks could no longer speculate on the stock market or lend on margin to stockbrokers and thus returned to the classic banking model, providing only short-term liquid loans to the commercial and agricultural sectors. Together with other changes, the financial sector as a whole, which was subject to an unprecedented level of federal regulatory oversight, had little incentive to take risks or to innovate, to such an extent that the period between 1940 and 1965 has been dubbed “boring banking.” However, the new regulatory apparatus was not without its weaknesses. Besides the failure to curb regulatory competition, the reforms of the 1930s created a stringent regulatory environment that prevented financial institutions from adapting to changes in macroeconomic conditions. The “boring banking” period, therefore, was simply the calm before a new financial storm.
3.5 Deregulation Amidst Regulatory Failures

The stable banking period that persisted for several decades following the Great Depression saw historically low bank failure rates. Nearing the tail end of the period leading up to the 1970s, a number of factors created challenges for the savings and loan (S&L) industry and banking sector. Financial innovations from the money market fund industry and various competitive pressures on S&L institutions led to disintermediation. In addition, the severe housing price declines in various regional markets, combined with limitations on inter-branch businesses and sector diversification, worsened the financial conditions of S&L institutions with asset portfolios concentrated in these regions.

Furthermore, there were tensions between deregulation in response to market innovation and safety and soundness supervision. The parallel regulatory structure and supervisory approach had weaknesses that fueled regulatory competition under the dual banking framework. Differences in regulatory policies among regulators and between federal and state regulatory entities engendered regulatory competition and complexity.

Amidst an unstable macroeconomic environment with rising interest rates, the financial industry relied a great deal on short-term savings deposits in order to fund long-term assets. These included higher interest rates associated with out-of-control inflation and tightened monetary policies. Higher interest rates also caused an asset-liability mismatch, resulting in widespread negative net worth and an insolvency crisis in the S&L and banking industries.

To tackle the immediate crisis, decision-makers instituted a package of measures that increased the social safety net for troubled firms, allowed regulatory
forbearance, relaxed investment activities, loosened barriers to entry, and bailed out large firms—prompting the notion of too-big-to-fail and reinforcing moral hazard, all of which contributed to higher bailout costs. This set of regulatory policies worsened financial conditions, and consequently exacerbated failures and costs to insurance funds.

Five Key Banking and Financial Laws

The five key laws that best represent the deregulation of the financial industry during this period are as follows.

First, the prevailing preference for deregulation led to the Depository Institutions Deregulation and Monetary Control Act of 1980 (DIDMCA), which removed interest rate control (Regulation Q). This substantial change is understood as a response to competitive conditions in the financial industry. However, the impetus for this legislation was planned years before its 1980 enactment. In 1972, a study by the Hunt Commission recommended phasing out interest rate caps on time and savings accounts as well as expanded lending and investment authority for financial firms. In 1976, Congress also issued a subcommittee report with recommendations along the same lines (Committee on Banking, Currency and Housing, 1976). When Congress passed the DIDMCA, it sought to offset competitive pressures from MMMFs by removing deposit interest rate caps. However, it had the unintended consequence of raising deposit funding costs and pushing many institutions to insolvency.

The second law Congress passed was the Garn-St. Germain Depository Institutions Act of 1982, which expanded the powers of saving associations to invest in nonresidential loans. It allowed the implementation of capital forbearance for thrifts facing the prospect of insolvency, while also relaxing portfolio
restrictions, which spurred an aggressive expansion into commercial real estate by thrifts. Banking problems attributed largely to real estate problems, and public policy responses made them even worse. This included tax laws that delayed taxable events and stimulated commercial real estate lending, bolstered by the Garn-St. Germain Act of 1982, which expanded the powers of saving associations to invest in nonresidential loans, combined with new entrants and heightened competition from commercial paper markets.

After a long hiatus, Congress enacted a third law: the Competitive Equality Banking Act of 1987 (CEBA). This initiative sought to address the depleting FSLIC, replenishing the fund with approximately $10.8 billion while allowing regulatory forbearance for certain thrifts and placing stricter standards of accounting and capital for the thrift industry as a whole. This bill started out as a part of an omnibus legislation that would expand the scope of commercial banks. However, due to Congressional fragmentation, the omnibus bill that was approved imposed a moratorium on banking regulatory agencies’ ability to grant new powers, which aimed to provide Congress with the opportunity to forge consensus, but to avail. When the moratorium expired, federal regulators at the OCC and the Fed took steps to expand the industry’s new powers.

Two focuses in financial regulation and supervision that ensure the safety and soundness of financial firms include the following: (1) microprudential—capital adequacy and liquidity and (2) market conduct—consumer protection, responsible lending, and governance. These were the prevailing focus of regulation and supervision during the S&L crisis. However, the government fell severely short on both of these aspects. Furthermore, a number of researchers
have noted that the government loosened financial rules amidst a wave of deregulated initiatives in order to enhance competitiveness without allocating sufficient examiners and supervisory resources to enforce the safety and soundness of institutions.

For policymakers, several key takeaways emerged from the S&L debacle. The first is that a regulatory structure is crucially important in banking supervision. From a supervisory perspective, the S&L crisis is a prime example of a supervisory structure with flaws that engendered a “race to the bottom” environment. Federal and state agencies were searching aggressively for new charter registrations, which created supervisory overlaps and incoherent enforcement of rules while allowing entities to benefit from regulatory arbitrage. In addition, regulated entities were insufficiently covered in that firms—specifically, S&L associations and commercial banks—were not uniformly subject to regulation and supervision. Economists and FDIC researchers also argued that deregulatory measures should have been offset by more bank examiners to cover shortages in light of higher-risk banks and stronger capital standards (Federal Deposit Insurance Corporation, 1997).

Second, the Federal Home Loan Bank Board’s (FHLBB’s) examination, supervision, and regulatory capabilities were not commensurate to the challenges when problem firms and failures surged. After the Great Depression, Congress established the Federal Home Loan System for the S&L industry, which was modeled after the Fed. With 12 regional Federal Home Loan Banks (FHLBs), the FHLBB oversaw the thrift industry in much the same way the Federal Reserve Board oversees commercial banking. However, the FHLBB’s structure had fundamental weaknesses and conflicts of interest. Regulatory reform sought to
address this structure; however, it did not sufficiently address the long-term problem of ill-equipped examination and supervisory capabilities. From this perspective, regulators and supervisors failed to achieve its two principal foci of microprudential concerns and market conduct.

Third, a major challenge for regulators during the S&L debacle was the need to balance regulatory discretion and principles-based regulation and supervision when dealing with distressed firms. During this period, regulators were supportive of discretion, which lent itself to regulatory forbearance. Many estimates point to costly results of such forbearance. Worse, forbearance had a tendency to favor large banks, thereby reinforcing too-big-to-fail market perceptions. Forbearance also included relaxing capital standards and accounting rules as well as allowing weaker institutions to either gamble out of insolvency or merge with others.

These considerations were the rationale for supervisory re-regulation. In particular, Congress approved the fourth law, the Financial Institutions Reform, Recovery, and Enforcement Act of 1989 (FIRREA). To address the then insolvent FSLIC—the FDIC’s counterpart for the thrift industry—the FIRREA established a new regulatory agency: the Resolution Trust Corporation. It was designed to resolve hundreds of insolvent thrifts, liquidate S&L assets (largely real estate), and use the proceeds to repay insurance to depositors. Moreover, Congress transferred the functions of the S&L insurance fund to the FDIC, dismantled the flawed thrift regulatory structure of the FHLBB, and shifted its legacy functions to the newly established Office of Thrift Supervision (OTS). As a result, the FHLBB was abolished, while the OTS absorbed these functions.
Meanwhile, the FSLIC was also abolished and the FLSIC Resolution Fund, administered by the FDIC, assumed all of its assets and liabilities. The OTS, in particular, received the authority to charter, regulate, examine, and supervise savings institutions. The FIRREA also restricted the lending activities of thrifts so that 70 percent of their assets had to be mortgage-related, such as residential, home equity, and mortgage-backed securities. Furthermore, it required thrifts to maintain capital standards comparable to those of national banks, with a minimum total capital ratio of 8 percent relative to assets.

In response to severe problems in the thrift industry and the nearly 1,300 commercial banks that either failed or required FDIC assistance over the two-decade span, Congress passed the fifth law—the Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA). With the goal of ensuring safety and soundness, the FDICIA placed restrictions on regulatory forbearance or intervention rules in the form of prompt corrective action and least-cost resolution. This required prudential regulators to take corrective supervisory actions in response to capital losses at a depository firm, based on categories corresponding to capital levels. It also required agencies to take corrective actions earlier and adopt the resolution method that minimized the cost to taxpayers. Finally, the law restricted institutions that were undercapitalized from borrowing at the Fed, on top of extending the examination cycle for banks and instituting a risk-based deposit insurance premium while increasing the FDIC’s credit line to the U.S. Treasury.
3.6 The 2007–09 Crisis and the Re-Reforms

3.6.1 The Rise of Shadow Banking

As the fast-changing macroeconomic conditions beginning in the 1970s challenged the stringent regulatory environment of the post-Great Depression era, a wave of deregulation washed over the financial industry. A host of regulatory barriers, which were installed in response to the Great Depression, were dismantled, making way for rapid financial innovations in all corners of the market. An unfortunate outcome of such deregulation was the rise of a new form of shadow banking, which contributed significantly to the fragility of the financial system.

As discussed in previous sections, shadow banking in the United States has had a long history, evolving from one form to the next. Before the Civil War, wildcat banks issued highly leveraged currency notes in remote rural areas and circulated them in large cities to minimize redemption. This was one of the earliest forms of shadow banking in America. At the turn of the 20th century, trust companies rose significantly in number and in asset size, competing directly with commercial banks while remaining unregulated (or lightly regulated). In the 1920s, investment trusts and investment affiliates made up the new shadow banking system, fueling rampant speculation in the stock market. Following the deregulation trend in the 1970s, it was natural that a new form of shadow banking would begin to take root.

Unlike previous forms, the modern shadow banking system is difficult to pinpoint, which is primarily because of the complexity of the modern financial system. According to the Financial Stability Board (2011), shadow banking
involves “credit intermediation involving entities and activities (fully or partially) outside the regular banking system.” Meanwhile, Claessens and Ratnovski (2014) describe it as “all financial activities, except traditional banking, which require a private or public backstop to operate.” By either definition, the shadow banking system includes a large number of different types of financial entities that specialize in complex credit-creation and funding techniques. Examples of these entities include: hedge funds, MMMFs, SIVs, SPEs, asset-backed commercial paper (ABCP) conduits, investment banks, and mortgage companies. Examples of credit-creation and funding techniques include securitization, repurchase agreements (repos), and risk-hedging through credit default swaps (CDS).

Financial deregulation contributed to the rise and the invasive nature of the modern shadow banking system in several aspects. First, financial barriers were removed in Congress. The deregulation trend, which began in the 1970s and culminated in the passage of the Gramm–Leach–Bliley Financial Services Modernization Act of 1999, dismantled the legal requirements separating investment and commercial banking. This bolstered the broader trend of consolidation in the financial services industry, thereby allowing a host of different types of financial institutions to operate under the same corporate roof. Meanwhile, the Commodity Futures Modernization Act of 2000 prevented the Commodity Futures Trading Commission (CFTC) from regulating credit derivatives and other over-the-counter derivatives, including CDS. Second, virtually all federal regulators joined the effort of deregulating the financial system via the exercise of their discretionary power. For instance, the Fed, which was responsible for the supervision of financial conglomerates, allowed these institutions to use their
bank affiliates to support non-bank loans, including shadow banking operations (Omarova, 2011). Meanwhile, the OCC modified its rules and regulatory interpretations so that banks under its supervision could engage in securitizing their own assets (Omarova, 2009). The Securities and Exchange Commission (SEC), which supervised investment banks, such as Bear Sterns, Lehman Brothers and Merrill Lynch, also decided to allow these banks to use their propriety risk models and set their own regulatory capital. As a result, their capital fell substantially, while their leverage increased (Office of Inspector General, 2008). Furthermore, the SEC exempted securitization vehicles from having to register as investment companies (Securities and Exchange Commission, 1992). This allowed these vehicles to continue operating in the shadows.

Amidst the new regulatory environment, the shadow banking system grew rapidly. According to Gorton and Metrick (2010), the assets of the shadow banking system were comparable to those of commercial banks in 2007. For instance, MMMFs, which provided short-term funding to the financial system at large, increased from $4 billion in asset size in the late 1970s to $1.8 trillion in 2000, peaking at $3.8 trillion in 2008. The repo market, which is a market for short-term collateralized loans, grew from $2 trillion in 1997 to $7 trillion in 2008. Furthermore, non-agency annual issuance of all securitized products, which was less than $100 billion in 1990, surged to more than $2 trillion just prior to the 2007–09 crisis.
3.6.2 It’s a Liquidity Crisis, Again!

In the run-up to the financial crisis, institutions were over-reliant on short-term unsecured funding markets to finance their operations. Specifically, financial firms relied on short-term money markets as a funding mechanism to support various asset-backed securitization businesses. The shadow banking system, which was heavily involved in securitization markets, amplified further the system’s vulnerability to liquidity shocks. For instance, the Lehman Brothers bankruptcy resulted in massive losses on commercial papers, causing Reserve Primary Fund (a MMMF) to “break the buck,” thereby causing investors to pull out in tandem. This sparked a general run on MMMFs and cascaded markets that relied on short-term funding. The rapid deterioration of financial market conditions prompted the Fed to create an array of special credit facilities to deal with liquidity pressures and credit shortages. The Fed’s actions can be summed up as providing massive liquidity and broadening the class of acceptable collateral, reflecting the complex web of intermediation activities performed by bank and non-bank firms.

When pressures emerged in term funding markets in August 2007, the Fed responded in several ways. First, it cut the discount window rate from 6.25 percent to 5.75 percent, which was 0.5 percentage points above the federal funds rate of 5.25 percent. Second, the Fed used its authority, granted under Section 10(b) of the Federal Reserve Act, to introduce the Term Auction Facility (TAF) in December 2007. Under this facility, the Fed provided term discount window loans in exchange for acceptable discount window collateral. This innovation allowed the TAF to avoid the stigma of the traditional discount window by
allocating the funds via auction to borrowers under sound financial conditions.

In March 2008, primary dealers faced liquidity pressures with respect to term funding access and collateral. Because broker-dealers had difficulties pledging collateral to finance their securities inventory, the Fed created the Securities Lending Facility (TSLF) on the basis of its Section 13(b) authority. Under this program, the Fed provided short-term liquid treasuries to primary dealers in exchange for a fee with collateral. In exchanging their low-quality collateral for high-quality collateral, primary dealers could pledge those in the repo lending market to raise funds. Similar to the TAF, the Fed auctioned off the TSLF funds, which peaked at $270 billion in October 2008.

During this period, the tri-party repo segment was also under strain, prompting the Fed to establish the Primary Dealer Credit Facility (PDCF). Due to their important role as counterparties to monetary policy implementation (by way of open market operations), the PDCF served as an overnight loan facility for primary dealers, in exchange for full collateral, which tracked the broader set of collateral accepted in the tri-party repo market. This credit facility mirrored that of the discount window extended to depository institutions. Meanwhile, outstanding credit peaked at $155 billion in September 2008.

Furthermore, in response to the credit crunch, the Fed, in coordination with the Treasury, took aggressive actions to restore the provision of credit lending in short-term lending markets. Under Section 13(3) of the Federal Reserve Act, the Fed created the ABCP MMMF Liquidity Facility to provide liquidity to depository institutions to fund the purchase of ABCP from MMMF's subject to considerable investor redemptions. The U.S. Treasury then instituted a temporary guarantee, with voluntary participation, of MMMFs by insuring their
shareholder assets up to the total value of shareholders’ accounts, which was important for MMMFs under redemption pressures. This guarantee expired in September 2008.

The Federal Reserve also established the Commercial Paper Funding Facility (CPFF). Under this facility, the FRB of New York purchased highly rated commercial paper at a three-month maturity directly from eligible issuers. In exchange, the Federal Reserve required issuers to pay fees to tap the facility, which proved to be a valuable alternative source of unsecured funds.

Disruptions in MMMFs erupted into a global liquidity crisis when firms retracted their dollar lending to one another due to concerns regarding their ability to fund operations, thereby resulting in shortages in the dollar-funding market overseas. When interbank money markets tightened considerably, unsecured funding sources eroded, and international firms were in need of dollars to fund their dollar-denominated assets, including securities holdings and retail and corporate loans. According to BIS economists, exposures were concentrated in Europe, where European banks had substantially increased dollar assets over time, with $8 trillion by mid-2007 (Baba, McCauley, and Ramaswamy, 2009). To address this shortage, the Fed entered into temporary central bank liquidity swap lines with a number of foreign central banks. While a number of foreign central banks were either unable or reluctant to tap into their foreign reserves supply, dollar availability from the Fed enabled these central banks to lend dollars to institutions in their jurisdictions.

In addition, to stop the financial system from hemorrhaging, the Fed leveraged massive publicly funded programs from the U.S. Treasury to maximize the impact of its credit facilities and financial assistance programs. Of note is the
expansion of the Term Asset-Backed Securities Loan Facility (TALF), which the Fed created in 2008 to support consumer lending. The TALF provides financing to private investors to help thaw the credit markets and lower interest rates for auto, small business, credit card, and other forms of consumer and business credit. Meanwhile, the Treasury committed $20 billion in “first-loss protection” from its Troubled Asset Relief Program (TARP) to leverage lending from the Federal Reserve. The goal was to increase the size of the TALF by using a $100 billion TARP commitment to leverage up to $1 trillion in new lending.

Invoking its “exigent circumstances” clause, the Fed worked with the Treasury to provide assistance to systemically important financial firms. This included aid to Citigroup, which had suffered substantial losses from investments in mortgage-backed securities as well as from credit rating downgrades. Under the agreement, the government would insure a pool of $306 billion in assets against unusually large losses, with Citigroup absorbing first losses. Any losses in excess of that amount would be shared between the government (90 percent) and Citigroup (10 percent). The government’s share of these losses would be allocated first to the Treasury through TARP, second to the FDIC, and finally, to the Fed through a non-recourse loan. The second notable rescue was AIG, which was facing imminent failure due to its inability to obtain sufficient capital or liquidity in private markets in order to function. The Treasury and the Fed jointly announced a package of measures to address the vulnerabilities in AIG’s balance sheet that threatened its viability as well as its credit rating. The FRB of New York borrowed billions in investment-grade, fixed-income securities from AIG in return for cash collateral. Meanwhile, the Treasury purchased preferred stock in AIG through TARP, with majority ownership, to provide additional
capital and help the company pay down its credit facility with the Fed.

From a bird’s-eye view, the 2007–09 crisis showed remarkably similar regulatory dynamics at play during the S&L crisis. This included the regulatory changes to accommodate private sector innovations, competitions between regulated sectors and less regulated sectors, and a complex regulatory structure with gaps. Just as in the S&L crisis, the recent financial crisis was attributed to various vulnerabilities. However, the most important distinction is that the 2007–09 crisis was a systemic liquidity crisis with global dimensions, while the S&L crisis was a fundamental asset-liability mismatch driving insolvencies. As a result, the regulatory reforms following each event were vastly contrasting.

3.6.3 Regulatory Re-Reform: The Dodd–Frank Act of 2010

After a year of public policy debate and almost two years after the financial system nearly collapsed, on July 21, 2010, Congress approved the Dodd–Frank Wall Street Reform and Consumer Protection Act, the most ambitious financial overhaul package since the Great Depression. By some measures, the aggregate word count in the Dodd–Frank Act that refers to financial restrictions dwarfs any legislation in modern history. Key provisions include (1) the creation of the Financial Stability Oversight Council (FSOC) and the Consumer Financial Protection Bureau (CFPB); (2) enhanced prudential oversight by the Fed and the FDIC; and (3) the regulation of systemic risk, consumer protection, sound underwriting standards, securitization risk-retention requirements, and over-the-counter derivatives clearing. Given the massive scope of the Dodd–Frank Act, the following section will focus more on the challenges related to liquidity
Money Market Mutual Funds

The Dodd–Frank Act charged the SEC with issuing reform rules to enhance the stability and resiliency of short-term funding markets. Under the FSOC’s direction, the SEC would recommend options to enhance the stability of MMMFs, including (1) increasing capital; (2) switching MMMFs from using a stable net asset value (NAV) to a floating NAV; and (3) implementing gates or liquidity fees to limit runs in times of stress.

However, the SEC came to a standstill because of internal disagreements. Due to overcapacity, complexity of rules, and disagreements over the risk trade-offs of the proposals, the SEC missed several of the original expected completion dates. Although the FSOC has asserted public pressure on the SEC to complete this ruling as well as an impact study, it does not have the authority to compel the SEC (or any other agency) to take specific regulatory or supervisory actions. In other words, the FSOC only serves as a coordinating body. This rules-issuance process highlights a weakness that officials have emphasized. For instance, Donald Kohn, former vice-chairman of the Fed, has suggested that the FSOC does not have sufficient independence and the power to impose requirements on individual regulator members (Fischer, 2014).

Given the uncertainty surrounding money-market funds, it remains to be seen whether the SEC’s reform rules will drastically reduce the risk of future panics/runs. This causes potential risks to financial stability. In addition, proposed changes, such as a floating NAV, could pose problems for treasury departments at financial firms, which would have to put systems in place to track fluctuations in NAVs as short-term capital gains and losses for tax purposes.
More broadly, with the regulatory changes, it is possible that money funds may become less attractive due to shifting demands from treasury departments and their cash portfolios.

**Systemic Risk: Tri-Party Repo**

Another market that shows financial stability concerns under distressed market conditions is the tri-party repurchase agreement system. This system relies heavily on the largest clearing banks, the Bank of New York Mellon and J.P. Morgan Chase, to facilitate the credit and collateral transfer between the lender and the borrower in a repo transaction. Given the nature of overnight repo transactions, until the clearing banks can match borrowers and lenders between a contract expiring and another contract beginning, clearing banks provide intraday (daylight) credit to borrowers to the tune of trillions of dollars. Therefore, tri-party repo participants are exposed to counterparty risks, including bank defaults and clearing bank defaults. Under a reform effort, the FRB of New York has pushed tri-party participants to reduce substantially their need for intraday (daylight) credit, down to as much as 20 percent. In a future scenario, such as a crisis facing the tri-party system, this would call for an active role by the Fed under its LOLR function to provide an alternative source of credit.

Beyond addressing intraday credit exposures, the tri-party repo market has not yet resolved the risk of collateral fire sales as a result of a liquidity crisis (Begalle, Martin, McAndrews, and McLaughlin, 2013). This continues to be a focal point for regulators and supervisors, given its important role in the broader repo lending markets. As witnessed during the crisis, primary dealers who borrowed as much as several trillion dollars a day in repo markets had difficulty
pledging low-quality collateral. Lenders who cannot hold the collateral on their books would be forced to sell in the event of a counterparty default. This could precipitate a momentum of margin calls, deleveraging, and asset price declines (Dudley, 2013). Systemic risk could also arise from either a defaulting dealer or a repo investor engaging in fire sales. Researchers have expressed concerns due to not only limitations of current tools to put controls on the risk of pre-default fire sales, but also the lack of tools to address post-default sales (Dudley, 2013).

Regulatory Gaps and Oversight Failure

As described, a major vulnerability observed pre-crisis was uncoordinated regulatory rulemaking. There was no single agency in charge of systemic risk. Furthermore, because of data gaps in monitoring excessive build-up of risk in the financial system, agencies need to leverage one another’s resources and expertise. From both microprudential and macroprudential perspectives, AIG is an illustrative example. It is one of the largest financial firms with a global footprint that opted for supervision under the OTS rather than the New York state banking regulator. The holding company relied on the solid ratings of its subsidiaries with traditional insurance products to issue CDS on senior tranches of collateralized debt obligations and other complex financial derivatives. Microprudential concerns shot to alarming levels when it posted substantial losses, a ratings downgrade, and significant collateral associated with CDS exposures. Consequently, the Dodd–Frank Act addressed these shortfalls by establishing the FSOC and by dissolving the OTS outright.

The FSOC is a unique solution to the regulatory architecture in the United States, an alternative to a single, consolidated regulator. The FSOC, which consists of voting members, including heads of federal banking and regulatory
agencies, serves as a policy forum to coordinate supervisory initiatives, identify, monitor, and address systemic risk. Furthermore, it has the authority to designate institutions as systemically important financial institutions that are subject to enhanced prudential standards by the Fed.

Like the S&L debacle, the sub-prime crisis revealed weaknesses in the regulatory structure that undermined the government’s ability to achieve both the traditional focus of microprudential regulation and market conduct. What compounded the impact of regulatory weaknesses was the link between financial stability and systemically significant financial firms that dealt with opaque, exotic products—all of which brought to the forefront another principal concern from a regulatory and supervisory perspective: regulatory failure due to competition in laxity. For instance, the OTS was perceived as a lax regulator vis-à-vis the FDIC. Even before the 2007–09 crisis, critics already voiced concern about financial firms taking advantage of the flexibility to choose charters. A number of them elected the OTS as their regulator based on the perception that the OTS was a lax regulator relative to alternative agencies, such as the New York State Banking Regulatory Agency. Examples include AIG, Washington Mutual, Countrywide, and IndyMac, all of which specialized in complex financial products before 2007 and then failed during the 2007–09 crisis under the OTS’s supervision. In response, the Dodd–Frank Act abolished the OTS agency and merged these functions into the OCC.
3.7 Conclusion

Via an examination of major regulatory reforms, this paper has provided a review of financial regulatory developments in the United States with special attention given to the issue of regulatory leakages. As reforms alter and shape the trajectory of the U.S. financial landscape over time, regulatory leakages, with respect to unregulated short-term funding markets, ineffective regulatory regimes, and shadow banking, evolve and manifest themselves in different forms. Thus, the discussion in this paper highlights the deep-rooted nature of these leakages.

The paper has also shown that misguided aspects of past reforms are among the key factors that contributed to market failures and ensuing panics. Because of this reciprocal and recurring relationship between regulatory overhauls and crises, it can be expected that factors leading to the next crisis may already exist in (or will emerge from) the current regulatory apparatus created by the Dodd–Frank Act. While the proximate cause is unknown, certain aspects of existing banking and financial regulation deserve a closer look from policymakers, lawmakers, and market participants alike. For instance, bank liquidity regulation should be updated frequently to account for financial innovations. In the recent crisis, the shadow banking system, which relied on repos and other short-term funding, was vulnerable to runs, just as depository institutions were before the Great Depression. Gorton (2010), Gorton and Metrick (2012), and others have described the recent market fallout as a classic bank run, driven by shadow banking firms’ inability to roll over short-term debt. Had the Dodd–Frank Act been implemented sooner, the 2007–09 financial crisis may have had a much
milder impact on the financial system and the economy as a whole.

Moreover, the unique regulatory regime and policy process in the United States should be acknowledged. While the regime includes a host of federal and state regulatory agencies as well as a web of complicated and sometimes overlapping rules and regulations, the public policy process is slow and reactive. Based on historical episodes and corresponding regulatory reform efforts highlighted in this paper, the optics of Congress may not allow for a comprehensive reform until a major crisis has come and passed. This is not an ideal situation. Perhaps, further examination of history will reveal more fundamental weaknesses that may ultimately convince Congress to take a more proactive position.

“Concisely stated, the problem before us at the present time is develop a sound banking system in the United States... why has the United States not found a satisfactory solution to its banking problems during the long period of agitation and reform that stretches back to 1836? Perhaps the chief reason is our habit of application—of being content with immediate remedies for pressing problems... To find a permanent solution we must first analyze the fundamental bases of the problem—not merely its superficial inconveniences.” — James (1934)
Bibliography


Appendix A

Appendix for Chapter 1

A.1 Two Supervisors and A Yardstick Competition

Let us consider the case that Assumption 1.24 does not hold. That means Congress can offer a supervisory contract that pays less when supervisors receive more banks. Then, to prevent supervisors taking advantage of this feature, I assume that supervisors cannot reject banks.

Consider a contract that pays:

\[ s(b_j, 0) = b_i \psi(e_C), \]

for \( i, j \in \{1, 2\} \).

This contract is similar to the yardstick competition contract proposed by Shleifer (1985). The main intuition is to pay a self-interested agent with non-verifiable effort not according to her own performance, but according to her peer’s. In this case, each supervisor receives transfer from Congress proportional to the number of banks that the other supervisor receives. Because bankers have the tendency to select a lower-effort supervisor (and because supervisors cannot reject banks), each supervisor will make sure that the other one has more banks.
by exerting the highest incentive-compatible level of effort. It follows that Congress can induce the supervisors to exert any $e_C$ that it desires.

In any case, it is not realistic to punish a supervisor for having too many banks. In particular, the form of supervisory contracts seen above is far from what can be observed in the real world.

A.2 Proof of Proposition 1.7.1

By imposing $K^* = 1 - \varepsilon$, bankers can borrow only $\varepsilon$ in deposits. Since all projects yield at least $\varepsilon$ in all states, bankers will always have enough assets to cover their debts. Hence, no bailout is needed in this case.

In addition, the interest rate on deposits is 0.

Besides deposits, bankers need to raise $B_{CoCos} = K^* - k$ in CoCos, at an interest rate $i_{CoCos} \geq 0$, for any level of equity capital $k \leq K^*$ they decide to contribute. Given supervisor effort $e$, bankers’ Incentive Compatibility Condition that ensures investments in safe projects is:

$$\bar{r}[\varepsilon - B_{CoCos}(1 + i_{CoCos}) - \varepsilon] \geq (1 - e)\bar{r}[R - B_{CoCos}(1 + i_{CoCos}) - \varepsilon]$$

$$\Leftrightarrow \bar{r}[\varepsilon - (K^* - k)(1 + i_{CoCos}) - \varepsilon] \geq (1 - e)\bar{r}[R - (K^* - k)(1 + i_{CoCos}) - \varepsilon]$$

$$\Leftrightarrow k \geq K^* - \bar{r}(IC(e) - \varepsilon),$$

where $IC(e)$ is given in (1.11). Note that the last inequality is obtained by recognizing that if bankers pursue safe projects with certainty, the interest rate
on CoCos satisfies:
\[ \bar{r}(1 + i_{CoCos}) = 1. \]

It is straightforward to verify that bankers’ profit is decreasing in \( k \). Hence, Condition (A.1) is binding. Given the total capital requirement \( K^* \) and supervisor effort \( e \), bankers will bring equity capital \( k(K^*, e) = K^* - \bar{r}(IC(e) - \varepsilon). \)

Furthermore, \( k(K^*, e) \) is decreasing in \( e \), because \( IC(e) \) is increasing in \( e \) as previously discussed. It follows that, similar to Section 1.6.1.4, Congress can implement any supervisory effort \( e_C \) by having two supervisors.

Congress’ welfare maximization problem becomes:
\[
\max_{e_C} W^{CoCos}(e_C) = w - b\psi(e_C) + b\left[ 1 - k(K^*, e_C) + \frac{\bar{r}}{\delta}(\varepsilon - (K^* - k(K^*, e_C)) - \varepsilon) \right]
\]

The first-order condition is:
\[ \psi'(e_C) = \bar{r}(1 - \frac{1}{\delta})IC''(e_C), \]
which is the same as Condition (1.21) in Section 1.6.2.3. It follows that, given \( K^* \), the optimal supervisory effort in this case is \( e_C = e^* \).

The required equity capital for bankers is then:
\[
k(K^*, e^*) = K^* - \bar{r}(IC(e^*) - \varepsilon) = 1 - \varepsilon - \bar{r}(IC(e^*) - \varepsilon) = 1 - (1 - \bar{r})\varepsilon - \bar{r}IC(e^*) < 1 - IC(e^*) = k^*,
\]
where the last inequality is obtained by recalling that \( IC(e^*) > \varepsilon \).
It is now straightforward to verify that

\[ W^{CoCos}(e^*) > W(e^*) \]

and \[ W^{CoCos}(e^*) > W^{Bailout}(k^*, e^*). \]

\[ \square \]

### A.3 Proof of Proposition 1.7.3

Note that from country G’s perspective, there are only two relevant choices for a capital requirement: 0 and \( k^0 \). That is because \( k^0 \) maximizes social welfare in a safe banking environment (by Proposition 1.6.4); whereas 0 capital requirement is optimal in a risky one.

**Case 1:** \[ \beta \frac{1-r}{1-\delta} \leq \alpha \leq 1 \]

For \( \beta \frac{1-r}{1-\delta} \neq \alpha \), \( F(k_L; \alpha, \beta) \) is strictly decreasing in \( k_L \). It follows that country L maximizes its welfare by setting \( k_L = 0 \).

If country G sets \( k_G > 0 \), its capital requirement is non-binding. All bankers register in L and enjoy 0 capital requirement. However, if \( k_G = 0 \), all bankers stay registered in country G. The share of bankers’ utility, \( F(0; \alpha, \beta) > 0 \), that would otherwise go to country L would remain with country G.

Thus, \( k_G = 0 \).

For \( \beta \frac{1-r}{1-\delta} = \alpha \), \( F(k_L; \alpha, \beta) \) is a positive constant with respect to \( k_L \). Hence, country L will always try to set \( k_L < k_G \) in order to attract all the banks, while country G maximizes its welfare by setting \( k_G = k_L \).

Thus, \( k_G = k_L = 0 \).

**Case 2:** \[ \beta \Omega < \alpha < \beta \frac{1-r}{1-\delta} \]
It is straightforward to verify that in this case, $\hat{k} < k^0$.

Also, $F(k_L; \alpha, \beta)$ is strictly increasing in $k_L$. It follows that country $L$ must set $k_L \geq \max\{0, \hat{k}\}$.

Country $L$ would not set $k_L \geq k^0$, since it would not get any banks, regardless of $k_G$.

Hence, $k^0 > k_L > \max\{0, \hat{k}\}$.

It follows that regardless of $k_G$, country $G$ is facing a risky banking sector. To maximize its welfare in this case, country $G$ must have $k_G = 0$.

Case 3: $0 \leq \alpha \leq \beta \Omega$

$F(k_L; \alpha, \beta)$ is increasing and $\hat{k} \geq k^0$. That means $F(k_L; \alpha, \beta) \geq 0$ only when $k_L \geq k^0$. Hence, country $L$ cannot lower its capital requirement to attract bankers.

Country $G$ maximizes its welfare the same way it does in the absence of country $L$: it sets $k_G = k^0$.

A.4 Motivation by Punishment?

Section 1.6.1 and 1.6.2 comprise the central analysis of this paper. They suggest that in a bailout-free environment, a supervisory structure should include multiple supervisors because they can be motivated to compete. However, in a bailout-guaranteed environment, it is difficult to motivate effective supervision, regardless of the number of supervisors.

These results depend on, among others, the assumption of non-verifiability of supervisory effort. Because of this assumption, supervisory contracts cannot
be directly contingent on supervisors’ effort. Instead, they are assumed to be contingent on the numbers of banks supervisors examine and close, \( \hat{b} \) and \( b^c \). In the time line of the model, both of these variables are revealed before Date 2.

Maintaining the assumption that supervisory effort is not directly verifiable, I relax assumptions on supervisory contracts, particularly allowing an additional transfer payment \( T \) from Congress to supervisors, which can be contingent on the number of failed banks, \( b^f \), at Date 2, for \( 0 \leq b^f \leq b \). The transfer payment, \( T(b^f) \), which can be negative, serves as yet another source of motivation for the supervisors. For instance, when \( T(b^f) > 0 \), it is a reward for a job well done; when \( T(b^f) < 0 \), it is a penalty. To ensure that a penalty is possible, I assume that supervisors receive a large endowment \( S > 0 \) at Date 2.

Before proceeding with the analysis, we first observe that the introduction of \( T(b^f) \) has no effect on the results in Section 1.6.1. That is because without a banking safety net, depositors and bankers always enter a deposit contract that ensures safe banking, regardless of supervisory effort exerted by supervisors. It follows that the number of failed banks, observable in Date 2, carries no new information regarding supervisors’ effort.

Thus, we will analyze the impact of \( T(b^f) \) on social welfare with the assumption that a bailout, as described in Section 1.6.2.1, is guaranteed to all failing banks.

Second, let us denote the probability of a crisis as \( 0 < \gamma < 1 \). In the initial description of the model in Section 1.4, I assumed that return shocks are perfectly correlated across all banks. Thus, if a crisis does not happen, all projects (safe and risky) succeed and yield high returns; if a crisis happens, a fraction of projects fail according to their types.
Specifically, when a crisis hits, safe projects fail with probability $\bar{r}'$ such that:

$$\bar{r} = (1 - \gamma) + \gamma \bar{r}'.$$ 

The ex-ante probability of failure of a safe project is, therefore, maintained at $\bar{r}$.

Similarly, when a crisis hits, risky projects fail with probability $r'$ such that:

$$r = (1 - \gamma) + \gamma r'.$$

Third, we observe that with a single supervisor, the number of failed banks, $b_f$, can take only three values: $b_f = 0$ if there is no crisis; $b_f = (1 - \bar{r}')b$ if a crisis occurs and all banks have taken on safe projects; and $b_f = (1 - r')b$ if a crisis occurs and all banks have taken on risky projects.

Respectively, let $T_0, \bar{T}$ and $\underline{T}$ denote the transfer payments a supervisor may receive when $b_f = 0$, $b_f = (1 - \bar{r}')b$ and $b_f = (1 - r')b$. I assume that $T_h \geq T_0, \bar{T}, \underline{T} \geq T_l$. The former inequality, which sets an upper limit for Date-2 rewards, reflects the fact that wages and bonuses in the public sector follow specific schedules and are generally capped by law. Meanwhile, the latter inequality, which sets the maximum penalty Congress can impose on supervisors, is a limited liability constraint that supervisors enjoy. In other words, supervisors are liable for bank failures only up to a certain extend.

### A.4.1 A Single Supervisor

Let us consider the case with only one supervisor.

First, recall that in Section 1.6.2.1, without supervision, a minimum capital requirement $k_0$ is necessary and sufficient to induce safe banking (Proposition 1.6.4). Social welfare is maximized at $W = W^k$. 

164
Second, recall that in Section 1.6.2.3, \( \bar{e}(k) \) given by (1.23) denotes the minimum level of supervisory effort that can ensure safe banking, for a given level of capital requirement \( k < k_0 \). Then, at Date 1-B, the supervisor faces the following scenarios when choosing effort \( e \):

- If \( e \geq \bar{e}(k) \), then \( \max_e U^s(e) = S + s(b) - b\psi(e) + (1 - \gamma)T_0 + \gamma \bar{T} \),
- If \( e < \bar{e}(k) \), then \( \max_e U^s(e) = S + s(b) - b\psi(e) + (1 - \gamma)T_0 + \gamma \bar{T} \).

In the first scenario, if the effort is high, the supervisor expects all banks to pursue safe projects. Thus, her expected utility depends on the difference between the transfer \( s(b) \) and the cost of supervision, and on the weighted average of the transfers she may receive when the number of failed banks is observed. The intuition for her expected utility in the second scenario is the same.

It is straightforward to see that it is optimal for the supervisor to choose either \( e = \bar{e}(k) \) or \( e = 0 \). The necessary condition for \( e = \bar{e}(k) \) is then:

\[
\bar{T} - T \geq \frac{1}{\gamma b\psi(\bar{e}(k))}.
\]  

(A.2)

Intuitively, Condition (A.2) means that the supervisor exerts high effort when either her potential Date-2 reward is sufficiently large, or her potential penalty is sufficiently severe.

Condition (A.2) and the following condition,

\[
S + s(b) - b\psi(e) + (1 - \gamma)T_0 + \gamma \bar{T} \geq S
\]  

(A.3)

are then necessary and sufficient to motivate the supervisor to exert \( \bar{e}(k) \), for a given level of \( k < k_0 \).

We have the following Lemma:
Lemma A.4.1. For each \( k < \bar{k}_0 \), there exists \( \{s(b), T_0, \bar{T}, T\} \) that satisfies Conditions (A.2) and (A.3) if and only if:

\[
T_h - T_t \geq \frac{1}{\gamma} b\psi(\bar{e}(k)).
\] (A.4)

Proof. First, for \( k < \bar{k}_0 \), consider that:

\[
T_h - T_t \geq \frac{1}{\gamma} b\psi(\bar{e}(k)).
\]

then, let \( T_0 = \bar{T} = T_h, T = T_t \), and

\[
s(b) = b\psi(\bar{e}(k)).
\]

It is straightforward to verify that Conditions (A.2) and (A.3) are satisfied.

Second, consider that:

\[
T_h - T_t < \frac{1}{\gamma} b\psi(\bar{e}(k))
\]

\[
\iff T - \bar{T} < \frac{1}{\gamma} b\psi(\bar{e}(k))
\]

which directly contradicts Condition (A.2). Note that the last inequality is obtained by observing that \( T_h \geq \bar{T}, T > T_t \).

At Date 1-A, Congress chooses a capital requirement \( k \) and a payment schedule \( \{s(b), T_0, \bar{T}, T\} \), to maximize:

\[
\max_{k, s(b), T_0, \bar{T}, T} W^{bf} = \quad w - \left[ s(b) + \bar{r}T_0 + (1 - \bar{r})\bar{T} + b(1 - \bar{r})(1 - k - \varepsilon) \right] \\
+ \left[ 1 - k + \bar{r} \left( \varepsilon - (1 - k) \right) \right] \\
+ S + s(b) + \bar{r}T_0 + (1 - \bar{r})\bar{T} - b\psi(\bar{e}(k)),
\]
subject to Conditions (A.2) and (A.3).

By Lemma A.4.1, the problem is simplified to:

$$\max_k W^{bf}(k) = w - b(1 - \bar{r})(1 - k - \varepsilon) + b \left[ 1 - k + \frac{\bar{r}}{\delta}(\varepsilon - (1 - k)) \right]$$

$$+ S - b\psi(\bar{e}(k)),$$

subject to Condition (A.4).

Let $\bar{k}$ be such that:

$$T_h - T_l = \frac{1}{\gamma} b\psi(\bar{e}(\bar{k})). \quad (A.5)$$

Solving the maximization problem, we have the optimal level of capital requirement is:

$$k^{bf} = \begin{cases} k_0 \text{ if } 0 \geq T_h - T_l, \\ \bar{k} \text{ if } \frac{1}{\gamma} b\psi(e^*) > T_h - T_l > 0, \\ k^* \text{ otherwise,} \end{cases}$$

where $k^*$ and $e^*$ are given by (1.21) and (1.22) in Section 1.6.2.2. Note that $\bar{e}(k^*) = e^*$.

Since $\bar{e}(k)$ is increasing in $k$ and $\psi(e)$ is increasing in $e$, it follows that $k^* \leq \bar{k} \leq k_0$. Correspondingly, $e^* \geq \bar{e}(k^{bf}) \geq 0$.

Hence, for a given level of $T_h - T_l$, Congress maximizes social welfare by implementing a capital requirement $k^{bf}$ and respective transfer payments that satisfy Condition (A.4) and motivate supervisory effort $\bar{e}(k^{bf})$.

It is then straightforward to verify that $W^{Bailout}(k^*, e^*) \geq W^{bf}(k^{bf}) \geq W^\bar{k}$.

In other words, Congress may be able to raise social welfare by motivating positive supervisory effort and lowering the capital requirement, pending the constraint on $T_h - T_l$. 

167
Note that Congress is able to reach $W^{\text{Bailout}}(k^*, e^*)$ only if $T_h - T_i$ is sufficiently large. From a policy making perspective, the threshold $\frac{1}{\gamma} b\psi(e^*)$, however, may pose a challenge to Congress because the cost of supervising the entire banking sector at $e^*$, $b\psi(e^*)$, may be large while the probability of a crisis, $\gamma$, can be very small.

### A.4.2 Two Supervisors And More

Given the analysis in the previous section, we now focus on whether Congress can improve social welfare by having two (or more) supervisors. As it turns out, having two (or more) supervisors can be socially desirable because the constraint on $T_h - T_i$ is substantially relaxed by the number of available supervisors.

The following proposition demonstrates the above result with two supervisors:

**Proposition A.4.2.** Let the minimum capital requirement $k^{bf}$ be such that:

$$k^{bf} = \begin{cases} 
  \bar{k} & \text{if } 0 \geq T_h - T_i, \\
  \frac{k^0}{k} & \text{if } \frac{b}{2\gamma}\psi(e^*) > T_h - T_i > 0, \\
  k^* & \text{otherwise},
\end{cases}$$

for $\bar{k}$ given by (A.5), and the contract transfer payments be:

\[
s\left(\frac{b}{2}, 0\right) = \frac{b}{2}\psi(\bar{e}(k^{bf})),
\]

\[
s\left(\frac{b}{2}, 0\right) \geq s(\hat{b}, 0), \forall 0 \leq \hat{b} \leq b,
\]

\[
T(b^f) = \begin{cases} 
  T_h & \forall b^f \leq (1 - \epsilon'')\frac{b}{2}, \\
  T_i & \text{otherwise.}
\end{cases}
\]

Then,

i) Each supervisors receives $\frac{b}{2}$ banks and exerts $\bar{e}(k^{bf})$,
\[ ii) \, k^0 \leq k^{bf} \leq k^*; \]

\[ iii) \, Social \, welfare \, is \, maximized \, at \, k^{bf} \, and \, W^k \leq W^{bf}(k^{bf}) \leq W^{Bailout}(k^*, e^*). \]

Proof. By design, we have:

\[ \gamma(T_h - T_l) \geq \frac{b}{2} \psi(\bar{e}(k^{bf})). \]

The supervisors’ incentive are as follow.

First, by exerting \( \bar{e}(k^{bf}) \), each supervisor will receive \( \frac{b}{2} \) banks, who will all invest in safe projects. Supervisors’ utility is:

\[
U^s_1(\bar{e}(k^{bf})) = U^s_2(\bar{e}(k^{bf})) = S + s(\frac{b}{2}, 0) - \frac{b}{2} \psi(\bar{e}(k^{bf})) + (1 - \gamma)T_h + \gamma T(1 - \bar{r}^{'}b)
\]

Second, suppose that supervisor \( j \) unilaterally chooses \( e < \bar{e}(k^{bf}) \). Due to the guaranteed bailout, all banks would prefer her. Supervisor \( j \)'s utility becomes:

\[
U^s_j(e) = S + s(\hat{b}, 0) - \hat{b} \psi(e) + (1 - \gamma)T(0) + \gamma T(1 - \bar{r}^{'}\hat{b}),
\]

\[
= S + s(\hat{b}, 0) - \hat{b} \psi(e) + (1 - \gamma)T_h + \gamma T_l,
\]

for any \( \hat{b} \geq \frac{b}{2} \) she decides to receive.

Since \( s(\frac{b}{2}, 0) \geq s(\hat{b}, 0), \forall \hat{b}, \) and

\[ \gamma(T_h - T_l) + \hat{b} \psi(e) \geq \frac{b}{2} \psi(\bar{e}(k^{bf})), \]

for \( \bar{e}(k^{bf}) > e \geq 0, \) it follows that \( U^s_j(\bar{e}(k^{bf})) \geq U^s_j(e), \) for all \( \bar{e}(k^{bf}) > e \geq 0. \)
Hence, supervisor $j$ has no incentive to deviate unilaterally from $e_1 = e_2 = \bar{e}(k^{bf})$.

Third, if both supervisors deviate, they would maximize their expected utility at $e_1 = e_2 = 0$. Similar to the analysis above, it is straightforward to verify that $\gamma(T_h - T_l) \geq \frac{b}{2} \psi(\bar{e}(k^{bf}))$ implies $U_{s}^{s}(\psi(\bar{e}(k^{bf}))) \geq U_{s}^{s}(0)$.

Therefore, $e_1 = e_2 = \psi(\bar{e}(k^{bf}))$ is a unique Nash equilibrium between the two supervisors. Each will then receive $\frac{b}{2}$ banks.

Now, the welfare maximization problem becomes analogous to the one in Section A.4.1.

Proposition A.4.2 shows that similar to Section A.4.1, Congress is able to motivate the supervisors to exert positive effort, and hence, achieve a high level of social welfare, pending the constraint on $T_h - T_l$.

The advantage of having multiple supervisors in this case is that the constraint on $T_h - T_l$ is substantially relaxed in comparison to the previous case. For instance, to achieve $W_{Bailout}^{Bailout}(k^{*}, e^{*})$, it is necessary in the case of a single supervisor that $\gamma(T_h - T_l) \geq b\psi(e^{*})$. However, in the case with two supervisors, Congress only needs the reward and penalty be such that $\gamma(T_h - T_l) \geq \frac{b}{2} \psi(e^{*})$.

From a policy making perspective, as Congress may face legal constraints on how much it can reward or penalize an individual supervisor, having multiple supervisors will make it easier for Congress to achieve its goals.
Appendix B

Appendix for Chapter 2

B.1 Proof of Proposition 2.6.1

Proof. First, it is straightforward that $k^* = \frac{1}{1-\rho}$ is necessary for a no-bailout commitment. That is because for $k > k^*$, the banker will issue risky deposits. In the event of a failure, the banker’s assets will be liquidated at a high cost. The government will, therefore, choose to bail out the bank.

Second, under the maximum leverage-ratio requirement $k^*$, for any investment size, $I$, the banker issues only risk-free deposits, $\rho I$. Hence, the bank will not be liquidated when it fails.

Now, given $k^* = \frac{1}{1-\rho}$, the banker solves:

$$\max_{B,r_B,I} U^b(B, r_B, I) = \alpha \left( \rho(I)I - (I + f - e_0 - B) - B(1 + r_B) \right)$$  \hspace{1cm} (B.1)

s.t. $\rho(I)I \geq I + f - e_0 - B + B(1 + r_B)$  \hspace{1cm} (B.2)

$$(e_0 + B - f)k^* \geq I$$  \hspace{1cm} (B.3)

$$\alpha(1 + r_B) \geq 1$$  \hspace{1cm} (B.4)
Solving the Kuhn–Tucker conditions, we therefore have:

\[ I = I^*, \]
\[ B = (1 - \rho)I^* + f - e_0, \]
\[ r_B = \frac{1 - \alpha}{\alpha}. \]

Since \( \eta > 0, [\rho(I^*)I^*] < [\rho(I)I'] \). Hence, \( I^* > \bar{I} \).

Social welfare, in this case, is:

\[ W^{CoCos} = w + \alpha \left( \rho(I^*)I^* - \rho \bar{I}^* \right) - f - (1 - \rho)I^* + e_0. \]

Furthermore, we have:

\[
W(e_0, \bar{I}) = w + \alpha \left( \rho(\bar{I}) \bar{I} - \bar{I} - f + e_0 \right) - (1 - \alpha)(1 + \eta)\left( (1 - \rho)\bar{I} + f - e_0 \right)
\]
\[
= w + \alpha \left( \rho(\bar{I}) \bar{I} - \bar{\rho} \bar{I} \right) - (1 - \rho)\bar{I} + e_0 - f - \eta(1 - \alpha)f
\]
\[
- \eta(1 - \alpha) \left( (1 - \rho)\bar{I} - e_0 \right)
\]
\[
< w + \alpha \left( \rho(\bar{I}) \bar{I} - \bar{\rho} \bar{I} \right) - (1 - \rho)\bar{I} + e_0 - f
\]
\[
< W^{CoCos},
\]

where the first inequality is due the banker’s limited equity and the second is due to the fact that \( I^* \) maximizes \( W^{CoCos} \).
Hai Xuan Nguyen was born in Hanoi, Vietnam, on June 21, 1985. He obtained a Bachelor of Science degree in mathematics and economics from the University of Tennessee, Knoxville (UTK) in 2007. Before entering the Ph.D. program in Economics at Johns Hopkins University (JHU) in 2008, he spent a year in Vietnam teaching mathematics and biking across the country. In 2010, he opted for a leave of absence from JHU to join the master’s program in mathematics at his alma mater, UTK. He earned his Master of Science in mathematics in 2011 and returned to JHU in the same year. As a Ph.D. candidate in the Department of Economics at JHU, Hai continues his research on banking and financial regulations. In 2012, he received the Dean’s Teaching Fellowship to develop a new course, “Financial Regulations in the U.S.: History and Framework,” which he taught independently in the fall of 2013 and spring of 2015. In August 2015, Hai will begin work as an assistant professor at the Chinese University of Hong Kong.