

Essays on Banking and Financial Intermediation

Jiayin Hu

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

©2019

Jiayin Hu

All Rights Reserved

ABSTRACT

Essays on Banking and Financial Intermediation

Jiayin Hu

I study financial intermediation and optimal regulation through the lens of banking theory and applied corporate finance. In my understanding, the theory on banking is primarily the theory on bank runs. And the key questions I have been pursuing to answer are the causes of runs in both the traditional and shadow banking sectors and the roles of the market and the regulator in maintaining financial stability.

I start with the shadow banking system outside the traditional regulatory framework, which accumulated tremendous risks and led to a major financial crisis. Why don't we simply shut down the shadow banking sector? **Chapter 1** examines the role of shadow banking and optimal shadow bank regulation by developing a bank run model featuring the tradeoff between financial innovation and systemic risk. In my model, the traditional banking sector is regulated such that it can credibly provide safe assets, while a shadow banking sector creates space for beneficial investment opportunities created by financial innovation but also provides regulatory arbitrage opportunities for non-innovative banks. Systemic risk arises from the negative externalities of asset liquidation in the shadow banking sector, which may lead to a self-fulfilling recession and costly government bailouts. Heavy regulatory punishment on systemically important shadow banks controls existing systemic risk and has a deterrent effect on its accumulation ex ante. My paper is the first to formalize the designation authority of a macro-prudential regulator in systemic risk regulation.

I then switch from the assets side to the liabilities side on the bank's balance sheet. **Chapter 2** introduces informed agents to the banking model and proposes a novel role of deposit insurance in fostering market discipline. While the moral hazard problem brought by deposit insurance weakens market discipline, I show that the opposite can

be true when the insurance stabilizes uninformed funding and increases the benefits of monitoring through information acquisition. Knowing the bank asset type, informed depositors utilize the demand deposits as a monitoring device and discipline the bank into holding good assets. However, self-fulfilling bank runs initiated by uninformed depositors erodes the future returns, inducing more depositors to forgo information acquisition and act like uninformed depositors. A novel role of deposit insurance emerges from the strategic complementarity between monitoring efforts and stability of uninformed funding. A capped deposit insurance, by stabilizing the retail funding of the bank, restores wholesale depositors' monitoring incentives and benefits market discipline.

I examine the role of information in generating bank runs in **Chapter 3**, where I explore the relationship between redemption price and run risks in a model of money market fund industry. Money market funds compete with commercial banks by issuing demandable shares with stable redemption price, transforming risky assets into money-like claims outside the traditional banking sector. Floating net asset value (NAV) is widely believed a solution to money market fund runs by removing the first-mover advantages. In a coordination game model *a la* Angeletos and Werning (2006), I show that the floating net asset value, which allows investors to redeem shares at market-based price rather than book value, may lead to more self-fulfilling runs. Compared to stable net asset value, which becomes informative only when the regime is abandoned, the floating net asset value acts as a public noisy signal, coordinating investors' behaviors and resulting in multiplicity. The destabilizing effect increases when investors' capacity of acquiring private information is constrained. The model implications are consistent with a surge in the conversion from prime to government institutional funds in 2016, when the floating net asset value requirement on the former is the centerpiece of the money market fund reform.

Contents

List of Tables	iv
List of Figures	v
1 Regulating Shadow Banks: Financial Innovation versus Systemic Risk	1
1.1 Introduction	2
1.2 The Baseline Model	12
1.2.1 The Environment	12
1.2.2 Banks and Financial Innovation	12
1.2.3 Laissez-faire Banking	14
1.3 Traditional and Shadow Banking	17
1.3.1 Imperfect Regulation	18
1.3.2 Coexistence Equilibrium	21
1.3.3 Welfare Comparison	22
1.3.4 Comparative Statics	22
1.4 Shadow Bank Regulation	23
1.4.1 Deterrent Effect of Designation	23
1.5 Systemic Risk and Credibility of Designation	27
1.5.1 Maturity Mismatch and Liquidity Risk	28
1.6 The Financial Stability Oversight Council (FSOC)	33

1.6.1	SIFI Designation	34
1.6.2	Enhanced Prudential Standards	36
1.6.3	Identifying Emerging Risks	37
1.7	Conclusions and Future Research	39
2	Bank Liability Structure and Optimal Deposit Insurance	41
2.1	Introduction	42
2.2	The Baseline Model	46
2.2.1	The Setup	46
2.2.2	Demand Deposits and Bank Runs	47
2.2.3	Benchmark: Equilibrium without Bank-asset Information	49
2.3	Equilibrium under Exogenous Information Acquisition	49
2.3.1	Equilibrium Analysis	54
2.4	Conclusion	55
3	To Float or not to Float? A Model of Money Market Fund Reform	56
3.1	Introduction	57
3.2	Background on the Money Market Fund Reform	60
3.3	The Baseline Model	62
3.3.1	The Setup	62
3.3.2	Money Market Funds	67
3.3.3	Equivalence between MMF Shares and Bank Deposits	70
3.3.3.1	Demand deposits	70
3.3.3.2	The Equivalence	73
3.4	A Model with Floating Redemption Price	73
3.4.1	Insured demand deposits	77
3.4.2	Secondary market trading	79
3.4.3	Stable NAV	81

3.4.4	Amortized-cost vs. Mark-to-market accounting	85
3.5	Conclusions	87
Bibliography		87
A Appendix of Chapter 1		93
A.1	Interbank Market	94
A.2	Enhanced Capital Requirement	95
B Appendix of Chapter 2		101
B.1	Model with a Wealthy Depositor	102
B.2	Model with Separated Contracts	103
B.3	A Deposit Insurance Fund facing Tradeoffs	109
B.3.1	The Environment	110
B.3.2	Equilibrium with Banks	113
B.3.3	Deposit Insurance	115
B.3.4	Continuous-State Extension	121
C Appendix of Chapter 3		127
C.1	First-best allocation	128

List of Tables

B.1	Consumption Patterns	115
B.2	Welfare	115
B.3	Parameter Values	120
B.4	Demand Deposit Contract	120
B.5	Parameter Values	126
B.6	Demand Deposit Contract	126

List of Figures

1.1	Coexistence Equilibrium of Traditional and Shadow Banking	20
1.2	Comparative Statics: θ	23
2.1	Strategies of Agents: Non-pivotal vs. Pivotal Mass of Informed Agents . . .	54

Acknowledgement

My first and foremost thanks with no doubt go to my committee members: Professors Patrick Bolton, Harrison Hong, Jose Scheinkman, Joe Stiglitz, and Jeffrey Gordon.

I am heavily indebted to my advisor, Patrick Bolton, for his encouragement and support throughout every stage of my research. Patrick has always been a great advisor. His influential work on contract theory, especially on the contingent allocation of control rights and its applications in various contexts, leads me to appreciate the succinctness and elegance of theoretical papers. While artfully presented in simple, concise models, the insights are also deeply rooted in reality, with an excellent understanding of institutional details while focusing on the factors of first-order importance. His guidance to me is both directly through meetings, talks, and emails and indirectly through his example as a professor, researcher, and scholar. Thanks to his tremendous support and help, I was able to overcome negative feelings such as self-doubts and frustration, and focus on what I could do to improve, survive, and thrive. I will continue working hard on my research and hope to live up to his expectations one day.

I am also profoundly grateful to Jose Scheinkman and Harrison Hong for their invaluable advice. They care about students and prepare us for the fierce competition of the job market and the academic career at large. I have always been impressed by Jose's sharpness and sense of humor. He keeps intense curiosity about the economy around the world and grasps the grit of new things immediately. Every time I meet with Jose, I will learn how to conduct research that can answer real questions and how to refine my points

and express them efficiently. Harrison's radical honesty is of utter importance to equip researchers with the necessary information and skills to distinguish intellectually stimulating research ideas and effort-worthy research directions. The experience of refereeing for the International Journal of Central Banking increases my exposure to the academic profession and adds a valuable perspective on evaluating solid research. Working as a teaching assistant to Harrison is truly rewarding and attending the classes is more like a bonus instead of an obligation.

I am very fortunate to have Joe Stiglitz on my committee, providing tremendous insights and inspiring perspectives on financial markets, regulation, and many more. I had the honor to meet Joe shortly after I arrived at Columbia five years ago. His profound contribution to economics, asymmetric information being only one aspect of it, enables us to stand on the shoulders of giants and inspires generations of economists. Joe is always willing to make time for students and genuinely cares about their future career and impact on the world at large, which he has demonstrated as a great role model. His sharp insights, vast curiosity, and witty comments have benefited generations of students and will keep going on.

I benefit a lot from Jeff Gordon's expertise in law and financial institutions. Jeff and Patrick have been working together for many years to teach the course Financial Crises and Regulatory Responses at Columbia, which I would recommend to anyone who would like to have a systematic and deep understanding on this topic. The teaching of Jeff and Patrick itself is intriguing and enlightening, a good exemplification of kindling a fire and sparking students' interests.

It has been a pleasant and rewarding experience participating in and serving as the student coordinator for the financial economics colloquium organized by Patrick, Jose, Harrison, and Matthieu Gomez, whose comments and questions are both constructive and helpful to the speakers and inspiring and intellectually stimulating to the audience. Not only have I received valuable advice when giving presentations, but I have benefited

a lot from attending the talks and reflecting on my research process.

I would also like to take the opportunity to thank Wei Jiang, Giorgia Piacentino, and Laura Veldkamp for their detailed advice and suggestions on my research and presentation skills. I thank Simona Abis, Charles Calomiris, Gregory Cox, Tri Vi Dang, Navin Kartik, Patricia Mosser, Suresh Naidu, Dan O’Flaherty, Tobias Salz, Tano Santos, Martin Uribe, Stijn Van Nieuwerburgh, Kairong Xiao, and participants at the Financial Economics Colloquium, Finance Free Lunch, and Becker Friedman Institute Macro-Financial Modelling Summer Session for their helpful feedback and suggestions.

As a Ph.D. candidate, I genuinely value the experience of being part of the scholarly community. For instance, doing mock interviews and cheering up each other with my fellow job market candidates: Chris Cotton, Yoon J. Jo, Xuan Li, Misaki Matsumura, Qiuying Qu, Anurag Singh, Danyan Zha, Qing Zhang, and Weijie Zhong. I have also received tremendous help from fellow students and Columbia alums both in and beyond my cohort: Adrien Alvero, Tuo Chen, Yi Cheng, Julong Feng, Tong Geng, Yifeng Guo, Chengcheng Jia, Ryan Kim, Mai Li, Haaris Mateen, Xueting Wang, Nancy Ran Xu, and Yue Yu. I am particularly grateful to Xiaodi Li, Yang Jiao, Xingye Wu, and Haotian Xiang for their absolute trust and support throughout the turbulent job market season. I thank Eric Chiang, Xiaodi’s husband, for taking the headshot for my website.

“If you love someone, send him/her to New York, for it is heaven. If you hate someone, send him/her to New York, for it is hell.” The same thing applies to the doctoral program. So it is a miracle for a Ph.D. to survive in the City of New York. Heaven or hell, the five years at Columbia has perhaps witnessed the best and worst versions of myself, and in a way prepared my spirits and will for future endeavors. If I can make it here, where else will I not succeed? It is an omen to which I feel obliged to respond, and as from *The Alchemist*, to realize one’s personal legend is one’s only real obligation.

To my parents and my younger brother

Chapter 1

Regulating Shadow Banks: Financial Innovation versus Systemic Risk

1.1 Introduction

Shadow banks played a central role in the 2007-09 financial crisis, the largest recession in the nearly eighty years since the Great Depression. Regulatory arbitrage is recognized as the main driver of the growth of the shadow banking system. Shadow banks do not face the same regulation as traditional banks, and the difference in regulation is the main driver of the growth in shadow banking. Shadow banks are able to circumvent traditional bank regulation because of how they are organized: Instead of taking deposits and making loans within a single entity, the shadow banking system issues non-deposit, money-like claims backed by financial assets created from a diverse range of loans and works through intermediation chains, in which shadow banks interact with one another and jointly perform the same liquidity, maturity and credit transformation roles as traditional banks. Enormous systemic risk thus accumulates outside the traditional regulatory framework¹.

Bringing shadow banks back under the regulatory umbrella is the focus of post-crisis reforms. Most notably, the Dodd-Frank Wall Street Reform and Consumer Protection Act (DFA) of 2010 established the Financial Stability Oversight Council (FSOC) and granted the council the statutory power to designate systemically important non-bank financial institutions and subject them to enhanced prudential standards. However, the effect of this unprecedented regulatory authority on the shadow banking system is understudied.

My paper fills this gap by developing a model to investigate the optimal design of shadow bank regulation under a tradeoff between financial innovation and systemic risk. In my model, the traditional banking sector is regulated such that banks are prohibited from excessive risk-taking and therefore credibly provide safe financial claims. By conducting financial innovation, innovative shadow banks create a superior category of risky assets and operate outside traditional regulation. However, non-innovative shadow

¹ See, for instance, [Pozsar et al. \(2010\)](#) and [Gorton and Metrick \(2010\)](#)

banks disguise themselves as innovative banks and contaminate the sector.

Moreover, the deterioration of asset quality may disrupt the interbank market and lead to a self-fulfilling crisis. The accumulation of systemic risk in the shadow banking sector causes a time-inconsistency problem for the regulator, who cannot commit not to extending the liquidity backstop to all shadow banks. The optimal regulation on shadow banks underscores the importance of reducing systemic risk and thus bailout costs and can be achieved through the deterrent effect of designation when the regulator has limited inspection capacity.

The baseline model consists of households, firms and banks. Firms need funding to invest in projects. Households are born with endowments, while banks are penniless. Patient households are indifferent between early and late consumption, while impatient households care only about early consumption. Banks intermediate between firms and households, producing diversified portfolios of individual firm projects and issuing asset-backed financial claims to raise funding from households ².

I introduce an endogenous financial innovation process to generate both bright and dark sides of risk taking. In my model, financial innovation is defined as the process of creating a new financial technology, which can be interpreted as research and development in the screening technology ³ rather than motivated by avoiding taxes and regulation ⁴. Banks that innovate are able to produce superior risky assets, which have higher expected returns than safe and inferior risky assets produced by the old technology. Banks, each of which is randomly matched with a household and unable to observe the preference types of the household, issue asset-backed financial claims to raise fund-

² To focus on the difference between traditional and shadow banking, I abstract from the complex shadow banking system and condense the long and intertwined intermediation chain into one entity, without loss of the essence of shadow banking as an unregulated sector.

³ For instance, [Laeven, Levine and Michalopoulos \(2015\)](#) provides a Schumpeterian model to characterize the dynamics of financial innovation, which helps screen entrepreneurs' capability to perform technological innovation and is therefore always beneficial to the economy.

⁴ The tax- or regulation-evasion motives for financial innovation are discussed in [Silber \(1983\)](#), [Miller \(1986\)](#), [Tufano \(1989\)](#), [Allen and Gale \(1994\)](#), [Merton \(1995\)](#), and more recently, in the form of securitization such as in [Gennaioli, Shleifer and Vishny \(2012\)](#) and [Gennaioli, Shleifer and Vishny \(2013\)](#).

ing from unsophisticated households, which cannot distinguish asset types. I capture the risk-shifting nature of debt contracts by assuming that given the interest rate, banks always invest in risky assets. Non-innovative banks take excessive risks and hold inferior risky assets, while the risk-taking of innovative banks is socially optimal. Banks with different innovation costs endogenously choose whether to conduct innovation.

I begin with a long-term financial contract with no maturity mismatch and no liquidity risk. Only patient households participate in banking, and impatient households will remain in autarky. Banks cannot commit to holding safe assets, and there is only one banking sector with risky assets. Non-innovative banks disguise themselves as innovative banks. The equilibrium depends on the fraction of banks with low innovation costs, which captures the degree of popularization of financial technology. When the mass of low-cost banks is small, they will be crowded out of the market due to a prohibitively high interest rate. As the mass increases, there will be a separating equilibrium in which only innovative banks participate in banking, while non-innovative banks cannot afford the interest rate, and this is eventually replaced by a pooling equilibrium in which non-innovative banks enter the market. This prediction is consistent with the anecdotal evidence/pattern that the deterioration of asset quality coincides with a market boom.

To add maturity mismatch and systemic risk into my model, I incorporate the financial innovation process in a bank run model to demonstrate how the non-innovative banks, despite representing a small fraction of the total, can destabilize the financial system. To attract impatient household, the banks offer demandable financial claims with both short- and long-term interest rates. Households can redeem the claims in either period 1 or period 2.

I propose a new strategic complementarity mechanism, whereby households run on the banks out of a fear of pecuniary externalities in the financial asset market. When facing early redemption requests, banks holding superior risky assets are able to borrow from sophisticated outside investors in the interbank market under a mark-to-market col-

lateral constraint, while banks holding inferior risky assets are not and thus have to file for bankruptcy and liquidate their assets, which has pecuniary externalities. Liquidation of inferior risky assets has an adverse effect on the underlying projects, which overlap with those composing superior risky assets, and as a result reduces the value of superior risky assets. Given other households' redemption behavior, patient households would demand early redemption if the total redemption exceeds the threshold value and would wait until the second period otherwise. Note that the early redemption request by the patient households will increase the amount of liquidation of inferior risky assets, which exacerbates the consequences for the economy. In anticipation of this development, sophisticated investors in the interbank market would refuse to lend to banks holding superior risky assets, leading to more asset liquidation. This negative feedback loop induces strategic complementarity among households and contributes to a self-fulfilling bank run, where the interbank market collapses, market liquidity evaporates, all assets are liquidated, and the economy is in a crisis state.

To capture the market friction from pecuniary externalities caused by inefficient asset liquidation, I examine government intervention through a regulator and "lender of last resort" (LOLR), which captures the key roles of a central bank. A regulator able to distinguish all types of assets (symmetric information) can implement the first-best allocation by imposing differentiated regulation based on the asset types. The traditional banking sector is regulated such that it can credibly provide safe assets, while the shadow banking sector is intentionally subject to lighter regulation to encourage investment in superior risky assets. For example, the regulation can be implemented through two sets of risk-based capital requirements: one on traditional banking, which imposes a high risk weight on all risky assets to prevent risk-taking activities, and the other on shadow banking, where the high risk weight only applies to inferior risky assets to prevent excessive risk-taking but allow for beneficial innovation.

The major difficulty in shadow bank regulation is distinguishing between innovative

banks and non-innovative banks, both of which hold risky assets. Recognizing the fact that the regulator is less informed than financial intermediaries regarding the nature of financial activities, I depart from the benchmark case by introducing asymmetric information. Under asymmetric information, the regulator is less informed than the banks and cannot distinguish between superior and inferior risky assets ex ante. The traditional banking sector is regulated such that it can credibly provide safe assets and has access to the government liquidity backstop (public liquidity), while an unregulated banking sector creates space for superior risky assets produced by financial innovation and borrows from the interbank market (market liquidity). However, When the regulator cannot distinguish between superior and inferior risky assets, bad banks holding inferior risky assets disguise themselves as good banks to circumvent existing regulation. The regulatory difference provides arbitrage opportunities for banks that take excessive risks, which contaminates the unregulated sector (shadow banking). The systemic risk causes a time-inconsistency problem for the regulator in providing public liquidity, which exacerbates excessive risk-taking in the shadow banking sector.

The regulator faces a tension between fostering financial innovation and reducing systemic risks. I propose a new approach to shadow bank regulation that highlights the deterrence effect. The regulation comprises two parts: inspection capacity and punishment severity. Each bad shadow bank has a probability of being designated as a systemically risky entity and facing the costly regulation and punishment. The regulator, which can only identify the true types of a fraction of banks, can increase the punishment on identified bad shadow banks to deter them from conducting regulatory arbitrage and taking excessive risks ex ante. Under effective deterrence, only good shadow banks operate in shadow banking, while banks that do not innovate choose to remain in traditional banking and hold safe assets.

My paper is the first to formalize the effect of the designation authority of the FSOC, established under Title I of the DFA of 2010, on systemically important unregulated non-

bank financial institutions in controlling risk accumulation in the shadow banking sector. The model demonstrates that given asymmetric information and limited inspection resources for shadow banks, heavy punishment is critical to deliver effective regulatory deterrence, which justifies burdensome regulation on the designated systemically important financial institutions (SIFIs).

The designation has a macro-prudential impact not only by controlling the existing systemic risks of the designated but also by deterring excessive risk-taking behavior by the undesignated. I propose that for a regulator with asymmetric information and limited resources, interim regulation is more desirable than ex ante regulation and ex post intervention. Instead of establishing a market entry requirement, the regulator requires information disclosure and constantly monitors the underlying risks. Rather than intervening after risks materialize, the regulator can take preemptive measures, such as SIFI designation, living wills and orderly liquidation authority, to curb the risk-taking activities of an intermediary to avoid it becoming “too systemic to fail”. The proactive regulatory measures not only reduce the systemic risks of a designated intermediary but also, and more important, deter the excessive risk-taking incentives of undesignated intermediaries. The deterrence effect increases the effectiveness of interim regulation, reducing systemic risk while fostering financial innovation.

Furthermore, I go beyond the specific designs of FSOC and examine the optimal regulation under information and enforcement constraints⁵, which is distinct from classical economic theory where the regulator is an omnipotent social planner with perfect information. In my model, a constrained regulator helps deliver the first-best through the deterrence power of supervision and designation. The effectiveness of deterrence depends on the regulator’s inspection capacity and punishment intensity, the latter of which should be large enough to remove shadow banks’ opportunistic mindset. Therefore an

⁵ First, the regulator is less informed than the market participants. Even when detailed information is disclosed per requirement, the regulator may not be able to assess the underlying risks of all financial activities due to limitations posed by manpower, time, expertise, and other resources.

over-punishment ex post can be desirable ex ante.

The key to solving the problem is to reduce systemic risks in the financial system before they materialize. The government should place greater emphasis on ex ante regulation and prevention than on ex post rescue and bailouts. Expanding deposit insurance to shadow banks will eliminate bank runs but also cause a severe moral hazard problem. An implicit government guarantee or expected bailout falls into the same category. The crisis has the symptoms of a liquidity dry-up, but the injection of public liquidity or the provision of a government guarantee during the crisis will exacerbate the moral hazard problem, contributing to the mispricing of risks in the market ex ante.

Empowering the regulator is essential for the effectiveness of regulation. Studying the optimality and effects of the FSOC is important for both academia and practitioners. It will advance the literature on financial stability and regulation and will offer an analytical framework to policymakers, highlighting the desirability of FSOC-style regulation in handling the tension between fostering financial innovation and maintaining stability.

The optimal design of the regulatory framework for the financial system also matters for monetary policy, since the migration of activities from the traditional to the shadow banking system will increase unregulated money creation and weaken the effects of monetary policies. A better understanding of financial regulation helps us to improve the effectiveness of monetary policy.

My paper contributes to several strands of literature on banking, financial intermediation and regulation.

Firstly, I develop a unified model to analyze the bright and dark sides of shadow banking. The rapidly growing literature on the role played by shadow banks in the run-up to the 2007-2009 financial crisis has greatly enhanced our understanding of the nature of the crisis and the dark sides of shadow banking. See, for instance, [Gorton and Metrick \(2012\)](#) and [Gennaioli, Shleifer and Vishny \(2013\)](#) on securitization and the repo market, [Acharya, Schnabl and Suarez \(2013\)](#) on ABCP conduits, and more recently, [Chernenko](#)

and Sunderam (2014), Schmidt, Timmermann and Wermers (2016), and Xiao (2018) on money market funds (MMFs) and Dang, Wang and Yao (2014), Wang et al. (2016), and Allen et al. (2017) on China's shadow banking system.

However, the shadow banking sector is modeled in these papers to be more risky and undesirable, while the regulator is assumed as a social planner who has the information, ability and willingness to implement optimal regulation, which may not be entirely true in the real world. Some papers start to view the shadow banking sector differently, recognizing the fact that the regulator is often less informed than the banks in terms of making investment decisions. The new tradeoff faced by the regulator is between preventing excess risk taking and giving the banks the flexibility to capture profitable investment opportunities. As shown in Ordonez (2018), the shadow banking sector can be welfare-improving since it provides "a channel to escape excessive regulation that is asymmetrically more valuable for banks with access to efficient investment opportunities". Ruan (2018) shows that shadow banking fills the gap of bank loan supply induced by regulation using data of surrogate intermediaries in China. Particularly, Ruan (2018) highlights how shadow banks could be beneficial complements to state-controlled traditional banking in making small- and medium-sized loans.

These findings on both the dark and bright sides of shadow banking provide useful ingredients for theoretical modeling. In my paper, the shadow banking sector accommodates both beneficial financial innovation and excessive risk-taking activities, consistent with this empirical evidence.

Furthermore, I incorporate systemic risk in the form of negative externalities of asset liquidation into a classical bank run model to characterize financial contagion and the rationale for government intervention. This brings together the classical banking literature on demandable financial claims and maturity mismatch problem, for instance, Bryant (1980), Diamond and Dybvig (1983), Jacklin (1987), Jacklin and Bhattacharya (1988), Alonso (1996), and Donaldson and Piacentino (2017), and the more recent financial contagion lit-

erature on fire sale and spillover effects such as [Chernenko and Sunderam \(2014\)](#) and [Schmidt, Timmermann and Wermers \(2016\)](#). With the realistic model settings, I go one step further and examine how regulation should be designed to mitigate the systemic risks generated by shadow banks, rather than taking the regulatory framework as given.

Secondly, I propose a new approach to optimal financial regulation, which contribute to studies on the expanded regulation on shadow banks. Shadow banks play dual roles in fostering financial innovation and accumulating systemic risk and the optimal shadow bank regulation is different from that on traditional bank. The topic is new, and the field is understudied. [Farhi and Tirole \(2017\)](#), where “special depositors and borrowers” are at the core of the analysis, proposes ring fencing between regulated and shadow banking. [Ordonez \(2018\)](#), which is the most closely related paper to mine, demonstrates that taxing the shadow banks and subsidizing regulated banks can implement the first-best allocation, since only those with superior risky assets will afford the tax and raise funding through shadow banking. However, it is a challenging task to determine the tax base and tax rates of various types of shadow banking activities. Moreover, the policy recommendation, which imposes a tax on efficient shadow banks to subsidize inefficient ones, seems counter-intuitive and places the former at a disadvantage.

As the FSOC has already been established for nearly a decade and has been exercising its supervision and designation authority since 2013, it is more relevant to examine the desirability of FSOC-style regulation on shadow banks. Particularly, the FSOC emphasizes nonbank financial companies that pose a threat to U.S. financial stability, targeting those that contribute most to systemic risk and government bailout expectations. This precise focus of regulation not only helps to prevent another financial crisis but also avoids burdening small- and medium-sized financial institutions.

Thirdly, I contribute to the literature on optimal financial regulation ⁶. The exist-

⁶ The paper also enriches the literature on general regulation, where (the threat of) punishment is widely used in areas of environmental protection and food security but rarely studied in the context of the financial industry. These areas may differ in terms of the unforeseeable financial risks and consequences (different

tence of shadow banks significantly weakens the effect of regulation on traditional banks, which can bypass regulation by moving activities off their balance sheets through shadow banking. Simply tightening existing regulation will induce more migration from traditional banks to shadow banks, as shown in [Plantin \(2014\)](#), [Huang \(2015\)](#) and [Begenau and Landvoigt \(2016\)](#). While these papers have done an excellent job in investigating the impact of shadow banks on the effectiveness of traditional bank regulation, they focus only on the dark side of shadow banking. They also fail to provide analysis of the potential of the regulator to expand the range of regulation and bring all financial activities under the same regulatory umbrella, which is exactly what the FSOC was created to accomplish. Instead of imposing an exogenous regulatory framework on bankers, I show how differentiated regulation emerges as an optimal response to the market. This explains why we see both regulated and unregulated financial intermediaries (usually playing the same credit, liquidity and maturity transformation roles) coexisting in equilibrium.

The remainder of the paper is organized as follows: Section 1.2 presents a new model with endogenous financial innovation and analyzes the optimal regulation under symmetric information. Section 1.3 examines the roles played by the traditional and shadow banks when the regulator is less informed than the banks but more informed than households. In Section 1.4, I examine the time-inconsistency problem faced by the regulator and demonstrate how the optimal shadow bank regulation reduces systemic risk and thus serves as a commitment tool against ex post bailouts. Section 1.6 discusses the practice of the FSOC in exercising its designation authority and offers policy recommendations. Section 1.7 concludes the paper.

from crime), but share many common points such as systemic risk, contagion, and government intervention.

1.2 The Baseline Model

1.2.1 The Environment

I begin with a baseline model where a unit-mass continuum of agents live for three periods, $t = 0, 1, 2$. Each agent is born in period 0 with an endowment of one unit wealth and consume in period 1 (early consumption c_1) and period 2 (late consumption c_2). The agents are risk-neutral and patient with a utility function $U(c_1, c_2) = c_1 + c_2$. Agents have access to the risk-free storage technology, which returns one-for-one every period.

There are firms in need of funding to finance long-term projects which take two periods to finish. However, the agents face prohibitively high costs of monitoring, verification and contract enforcement and thus cannot directly provide funding to the firms. In other words, the direct financial market is not viable. In autarky, households store their wealth in period 0 and consume their wealth in period 1 and period 2, $c_2 = 1 - c_1$.

1.2.2 Banks and Financial Innovation

I motivate the role of financial intermediaries, which I refer to as "banks", by their advantage over agents in monitoring firms. Banks make loans to firms and raise funding from agents through financial claims. Specifically, I assume the banks can produce two types of assets out of firm loans: safe (type- f) and inferior risky (type- i). The payoff of safe assets, R_f , is a constant greater than one⁷. The inferior risky assets default and become worthless with probability q and succeed and return R_i with probability $1 - q$.

In addition to the traditional assets, banks may obtain access to a superior category of risky assets (type- s) through financial innovation (for instance, research and development in the financial industry to improve screening and risk management skills). Specifically, I assume that the risky assets have the same probability of default, but superior risky assets

⁷ Safe assets include but are not limited to U.S. treasuries and securities backed by them. Loans to creditworthy clients with sufficient, quality collateral are also considered "safe". The business model has matured: e.g., abundant historical data to trace back for analysis, time-proven methods to screen borrowers and manage risks, state-of-the-art models to evaluate collateral.

maintain a positive residual value $R_a \in (0, 1)$ upon default and return $R_s > R_i$ in period 2 when successful.

Assumption 1 (Payoffs of Assets). *The payoffs of assets satisfy the following assumptions:*

1. $R_s > R_i$: *The payoffs of superior risky assets strictly dominate those of inferior risky ones;*
2. $(1 - q)R_s + qR_a > R_f$: *The expected payoff of superior risky assets is higher than that of safe ones;*
3. $(1 - q)R_i < R_f$: *The expected payoff of inferior risky assets is lower than that of safe ones.*

I assume that the probability of project failure is neither too high nor too low, $q \in (\underline{q}, \bar{q})$, where $\underline{q} = 1 - \frac{1}{R_i}$ and $\bar{q} = 1 - \frac{R_f - 1}{R_i - 1}$,⁸

Financial innovation is not a free lunch. It incurs an upfront cost $\chi \in (0, 1)$ on a bank that conducts innovation and banks differ in innovation costs. The innovation is socially optimal only when its benefits (expected payoffs) outweigh its costs (innovation cost plus opportunity cost). Prior to innovation, safe assets are preferred to risky assets, therefore the socially optimal innovation strategy is characterized by χ^* , where

$$\chi^* = 1 - \frac{R_f}{(1 - q)R_s + qR_a} \quad (1.2.1)$$

Since it can observe each bank's innovation cost and asset type, the social planner will only allow banks with innovation costs lower than the threshold value χ^* to innovate and hold superior risky assets, while force the rest of the banks to produce safe assets. In the first-best allocation, two banking sectors coexist: a *traditional banking* sector composed of banks that hold safe loans and an *innovative banking* sector composed of banks holding

⁸ When $q > \bar{q}$, safe assets are preferred to superior risky assets, and the financial innovation technology becomes worthless. If $q < \underline{q}$, inferior risky assets have positive expected net value and may even generate higher yields in expectation than safe assets, which makes risk-taking desirable all the time, and the analysis of excessive risk-taking no longer matters.

superior risky loans. Both sectors are protected by the liquidity backstop. The coexistence of traditional and innovative banking encourages efficient financial innovation and prevents the creation of inferior risky assets.

Note that the threshold value of the socially optimal innovation strategy depends solely on the payoffs of safe and superior risky assets. I denote $\theta(\tilde{\chi})$ as the fraction of banks with innovation costs lower than a threshold value $\tilde{\chi}$. Given $\tilde{\chi}$, the corresponding value of θ measures the level of popularization of frontier technologies. With the advancement of technology, θ increases and more banks adopt the low-cost innovation technology, which facilitates the financing of superior risky assets.

1.2.3 Laissez-faire Banking

We now turn to the market equilibrium, where prices (interest rates) replace the social planner's commands and agents cannot observe the banks' innovation costs and asset holdings. I begin with laissez-faire banking to demonstrate market frictions and then analyze how government intervention improves social welfare.

Timeline. At the beginning of period 0, each agent is randomly assigned a bank holding assets and issuing financial claims to raise funding from the agents. Banks choose whether to innovate and which type of firm to make loans to and propose a take-it-or-leave-it offer to raise funding from households. Households choose between remaining in autarky and participating in banking. In period 2, projects are finished, and asset payoffs realize. Households redeem their claims and consume. Banks retain the residual value of the proceeds.

Although the inferior risky assets are not socially desirable, the risk-shifting nature of the debt contract may induce banks to take excessive risks:

Assumption 2 (Excessive Risk-taking Incentives). *Given interest rate r ,*

1. $(1 - q)(R_i - r) > R_f - r$: *Banks that do not innovate prefer inferior risky assets to safe assets;*

2. $(1 - q)[(1 - \chi)R_s - r] > (1 - q)[(1 - \chi)R_i - r]$: Banks that innovate prefer superior risky assets to inferior risky assets.

Assumption 2 excludes the possibility of a credible commitment by unregulated banks. Non-innovative banks have incentives to deviate from holding safe assets to holding inferior risky assets, of which the expected payoff falls below the original endowment. The excessive risk-taking incentive constrains banks' ability to issue safe claims, and unsophisticated households lack effective tools to discipline banks' risk taking behaviors.

Banks choose to innovate only if the innovation cost is small enough to make innovation more profitable than holding inferior risky assets, $(1 - q)[(1 - \chi)R_s - r] > (1 - q)(R_i - r)$, namely when

$$\chi < 1 - \frac{R_i}{R_s} \equiv \hat{\chi} \quad (1.2.2)$$

Here, the innovation strategy depends on the payoffs of risky assets upon success. This is because the banks make profits on the upside of the assets while the agents receive all the residual values on the downside. When R_i increases, the threshold value $\hat{\chi}$ decreases, meaning that more banks will be lured into holding inferior risky assets instead of creating superior assets. The model captures the deterioration of asset quality during economic booms⁹.

Generally, $\hat{\chi}$ is different from χ^* . I assume $\hat{\chi} < \chi^*$. In this case, the banks with $\chi \in (\hat{\chi}, \chi^*)$, which should have innovated in the first-best allocation, choose not to do so because holding inferior risky assets is more lucrative than investing in R&D. The fraction of innovative banks shrinks from $\theta(\chi^*)$ to $\theta(\hat{\chi})$. This is because non-innovative banks take

⁹ When different probability of success is assumed (q_s for superior risky assets and q_i for inferior risky ones), then the threshold value of innovation strategy becomes

$$\chi < 1 - \frac{\frac{1-q_i}{1-q_s}(R_i - r) + r}{R_s} \equiv \hat{\chi}'$$

The analysis above still applies and the new threshold value becomes smaller when the probability of default of inferior risky assets $(1 - q_i)$ decreases. Hence an additional implication emerges: during economic booms, where the default risk is small, more banks will prefer inferior risky assets to their superior counterpart which requires innovation costs.

advantage of the asymmetric information and disguises themselves as innovative banks. The pooling of superior and inferior risky assets increases the financing cost of innovative banks, crowding out good banks which should innovate.

Since they cannot distinguish assets, the agents make decisions based on their beliefs on the pool of innovative and non-innovative banks, which face the same interest rate. The interest rate offered by the banks makes the agents indifferent between participate in banking and storing their wealth,

$$r_{LF} = \frac{1 - q\theta(\hat{\chi})R_a}{1 - q} \quad (1.2.3)$$

In laissez-faire banking equilibrium, a fraction of $\theta(\hat{\chi})$ banks conduct financial innovation and gain access to superior risky assets, while the rest of the banks remain non-innovative and hold inferior risky assets.

The social welfare under laissez-faire banking is $SW_{LZ} = \theta(1 - \chi)[(1 - q)R_s + qR_a] + (1 - \theta)(1 - q)R_i$. There are several inefficiencies in laissez-faire banking: Firstly, excessive risk taking by non-innovative banks raises the funding cost of innovative banks, which would in some cases crowd out the latter. Secondly, innovation incentives are distorted since low-cost banks are tempted to create inferior risky assets. Thirdly, an excessively large shadow banking sector reduces the desirability of the overall banking sector, which is plagued by excessive risks and may no longer be viable. Due to the lack of commitment in creating safe assets, financial intermediation may break down and the benefits of banking have to be forgone.

The banking sector is viable only when the fraction of innovative banks is large enough:

$$\theta > \frac{1 - (1 - q)R_i}{(1 - \chi)[(1 - q)R_s + qR_a] - (1 - q)R_i} \quad (1.2.4)$$

1.3 Traditional and Shadow Banking

A regulator with symmetric information restores the first-best allocation from laissez-faire banking by imposing regulation based on the types of asset holdings. The regulation is as follows: (1) Only f -banks are permitted in the traditional banking sector, where banks credibly produce safe short-term financial claims; (2) only s -banks are allowed to operate in the innovative banking sector; and (3) i -banks will not be granted banking licenses. One example of such regulation is risk-based capital requirements, as illustrated in ? ¹⁰.

Note that the regulation is lighter in the innovative banking sector than in the traditional banking sector in the sense that banks are allowed to hold risky assets. When the regulator has perfect information regarding the true types of bank asset holdings, the difference in regulation creates space for socially desirable risk-taking by allowing innovative banks to operate. However, when the regulator cannot distinguish between superior and inferior risky assets ex ante, the license-granting regulation in the innovative banking sector is ineffective creates regulatory arbitrage opportunities ¹¹ for non-innovative banks, which, under symmetric information, can only survive by holding safe assets in the traditional banking sector.

¹⁰ Suppose that each bank is endowed with κ , which cannot be used for investment but can be seized when the bank fails. The regulator can impose risk-based capital requirements on banks to prevent excessive risk taking. The regulation consists of a capital requirement (minimum capital-to-asset ratio) Ψ and a vector of risk weights on assets $\omega = (\omega^f, \omega^s, \omega^i)'$. The regulator requires a bank's risk-adjusted capital-to-asset ratio to be higher than the threshold value. To foster financial innovation while preventing excessive risk taking, the regulator adopts different capital requirements, Ψ_{TB} and Ψ_{SB} , on traditional and shadow banks, respectively. Given the risk weights, an effective capital requirement on the traditional banking sector is satisfied only by f -banks, while an effective capital requirement on the shadow banking sector is satisfied by both f -banks and s -banks:

$$\frac{\kappa}{\omega^s V^s} < \Psi_{SB} < \frac{\kappa}{\omega^s V^s} < \Psi_{TB} < \frac{\kappa}{\omega^f V^f}$$

¹¹ For instance, commercial banks or bank holding companies establish special purpose vehicles, or conduits, to move lending activities off their balance sheets. The conduits issue asset-backed commercial paper or repos to obtain wholesale funding in the money market, rather than demand deposits. Another example is when non-bank financial institutions, such as investment banks, insurance companies and finance companies, engage in such activities.

1.3.1 Imperfect Regulation

Assumption 3 (A Regulator with Imperfect Information). *The regulator is able to distinguish between safe and risky assets but not between superior and inferior risky assets in period 0.*

The assumption captures the reality that the regulator has superior information collection ability than common households but is less informed than the banks. Under asymmetric information, *i*-banks disguise themselves as *s*-banks and apply for a license in innovative banking. Thus, the innovative banking sector is contaminated and becomes a shadow banking sector accommodating both innovative and non-innovative banks.

Still, the traditional banking sector improves welfare by providing a commitment device for banks to hold safe assets. The regulator only grants *f*-banks license to operate in the traditional banking sector. The regulation can be enforced perfectly since the regulator can distinguish between safe and risky assets.

However, the innovative banking sector now becomes a shadow banking sector which non-innovative banks can use to circumvent traditional bank regulation and take excessive risks. It creates space for good shadow banks (*s*-banks) that innovate and create superior risky assets *and* providing regulatory arbitrage opportunities for bad shadow banks (*i*-banks) that do not innovate and produce inferior risky assets.

The market equilibrium is characterized as follows: Banks follow sequential strategies. The first is an innovation strategy: Conditional on the bank's innovation cost, it chooses whether to conduct financial innovation, that is, $\mathbb{1}_{\mathcal{A}}: \chi \rightarrow \{1,0\}$, where \mathcal{A} denotes the choice of conducting financial innovation. The second is a financing strategy: Conditional on the bank's access to assets, it chooses the probability σ_F of using shadow banking to raise funds. The third is an asset holding strategy: Conditional on the innovation strategy and the regulatory environment, the bank chooses the type of assets k to make loans to \mathcal{K} : $\mathbb{1}_{\mathcal{A}} \rightarrow \{s, i, f\}$. The participation constraint of households is satisfied.

Banks prefer shadow banking to traditional banking when the expected profit of the

former is larger than that of the latter: namely $(1 - q)[(1 - \chi)R_s - r_{SB}] > R_f - r_{TB}$ for innovative banks and $(1 - q)(R_i - r_{SB}) > R_f - r_{TB}$ for non-innovative banks. Hence we have

$$\text{Innovative banks: } r_{SB} < \frac{1}{1 - q}r_{TB} + (1 - \chi)R_s - \frac{R_f}{1 - q} \quad (1.3.1)$$

$$\text{Non-innovative banks: } r_{SB} < \frac{1}{1 - q}r_{TB} + R_i - \frac{R_f}{1 - q} \quad (1.3.2)$$

Denote σ_l the probability household believes that low-cost banks will raise funding in shadow banking and σ_h the probability of high-cost banks participating in shadow banking. Therefore the fraction of shadow banks hold superior risky assets is given by

$$\sigma = \frac{\theta\sigma_l}{\theta\sigma_l + (1 - \theta)\sigma_h} \quad (1.3.3)$$

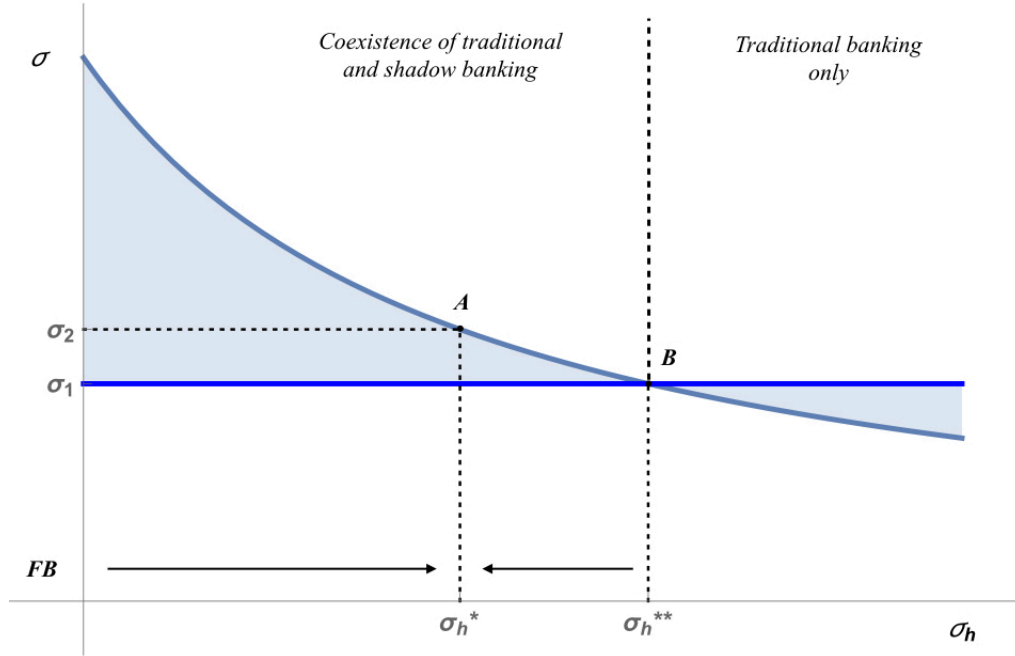
I focus on equilibria in which these beliefs are correct, as in [Ordenez \(2018\)](#). Given interest rate r_{TB} , the participation constraint of households is

$$r_{SB} = \frac{1 - \sigma q(1 - \chi)R_a}{1 - q} \quad (1.3.4)$$

Hence, given the agents' beliefs, banks adopt the following threshold strategy: low-cost banks will participate in shadow banking if $\sigma > \sigma_1$ and participate in traditional banking otherwise, while high-cost banks participate in shadow banking if $\sigma > \sigma_2$ and participate in traditional banking otherwise, where $\sigma_1 = \frac{[R_f - (1 - q)(1 - \chi_l)R_s]}{q(1 - \chi_l)R_a}$, $\sigma_2 = \frac{[R_f - (1 - q)R_i]}{q(1 - \chi_l)R_a}$, and $\sigma_1 < \sigma_2$.

There are three main implications of the equilibrium: First, the first-best allocation, represented by $(\sigma_l, \sigma_h) = (1, 0)$, cannot be sustained as an equilibrium. When $\sigma = 1$, the interest rate in shadow banking is so low that high-cost banks will make more profits by migrating their activities to the shadow banking sector. The deviation of high-cost banks

Figure 1.1: Coexistence Equilibrium of Traditional and Shadow Banking



reduces σ , but as long as $\sigma > \sigma_2$, the high-cost bank will have incentives to deviate from traditional banking to shadow banking.

Second, there is an equilibrium in which all banks remain in traditional banking. When $\sigma < \sigma_1$, even the low-cost banks cannot afford the high interest rates in shadow banking. The decrease in σ_l triggers a downward spiral, where a lower σ reinforces the incentives to leave shadow banking. In equilibrium, $\sigma_l = \sigma_h = 0$.

Third, there is an equilibrium in which traditional and shadow banks coexist. Low-cost banks will participate in shadow banking as long as $\sigma > \sigma_1$. When $\sigma \in (\sigma_1, \sigma_2)$, high-cost banks will shift away from shadow banking, which decreases σ and eventually brings σ back to σ_2 . When $\sigma > \sigma_2$, high-cost banks in traditional banking will be attracted to shadow banking, which increases σ_h and reduces σ . Thus, in equilibrium, it must be that $\sigma = \sigma_2$.

1.3.2 Coexistence Equilibrium

Let us now focus on the case in which innovative and non-innovative banks coexist in the shadow banking sector, namely when the quality of risky assets deteriorates.

At $(\sigma_l, \sigma_h) = (1, \sigma_h^*)$, all innovative banks will participate in shadow banking, while high-cost banks will participate in shadow banking with a probability σ_h^* . Given $\sigma_h = \sigma_h^*$, no low-cost banks will deviate from the strategy $\sigma_l = 1$, since their expected profit in shadow banking is higher than that in traditional banking. Given $\sigma_l = 1$, no high-cost banks will deviate from the strategy $\sigma_h = \sigma_h^*$. A higher σ_h decreases σ and makes traditional banking more profitable, which will reduce σ_h . Similarly, a lower σ_h increases σ and makes shadow banking more lucrative, which will bring σ_h back to the equilibrium level.

In the coexistence equilibrium, the size of shadow banking is $\sigma_{SB} = \theta + (1 - \theta)\sigma_h^*$, while the size of traditional banking is $\sigma_{TB} = (1 - \theta)(1 - \sigma_h^*)$. Due to regulatory arbitrage, the size of shadow banking is larger than the first-best.

Note that in the coexistence equilibrium, the participation constraint of high-cost banks is binding, while that of low-cost banks is slack. The interest rate in the coexistence equilibrium thus makes marginal non-innovative banks indifferent between participating in traditional or shadow banking:

$$r_{SB} = \frac{1}{1 - q} r_{TB} - \left(\frac{R_f}{1 - q} - R_i \right) \quad (1.3.5)$$

According to households' participation constraint, the interest rate should satisfy:

$$r_{SB} = \frac{1}{1 - q} r_{TB} - \frac{q}{1 - q} \frac{\theta}{\theta + (1 - \theta)\sigma_h} R_a \quad (1.3.6)$$

Combining equation 1.3.5 and 1.3.6, we obtain the equilibrium σ_h^* :

$$\sigma_h^* = \frac{\theta}{1 - \theta} \left[\frac{q(1 - \chi)R_a}{R_f - (1 - q)R_i} - 1 \right] \quad (1.3.7)$$

1.3.3 Welfare Comparison

Given the equilibrium (σ_l^*, σ_h^*) , the welfare of the coexistence equilibrium is given by

$$SW_{TBSB} = (1 - \theta)(1 - \sigma_h^*)R_f + \theta(1 - \chi)[(1 - q)R_s + qR_a] + (1 - \theta)\sigma_h^*(1 - q)R_i \quad (1.3.8)$$

The shadow banking sector is welfare-improving compared to the case where only traditional banking is permitted if $SW_{TBSB} > SW_{TB}$. That is to say, completely shutting down the shadow banking sector is not socially desirable when

$$\sigma_h^* < \frac{\theta}{1 - \theta} \frac{(1 - \chi)[(1 - q)R_s + qR_a] - R_f}{R_f - (1 - q)R_i} \quad (1.3.9)$$

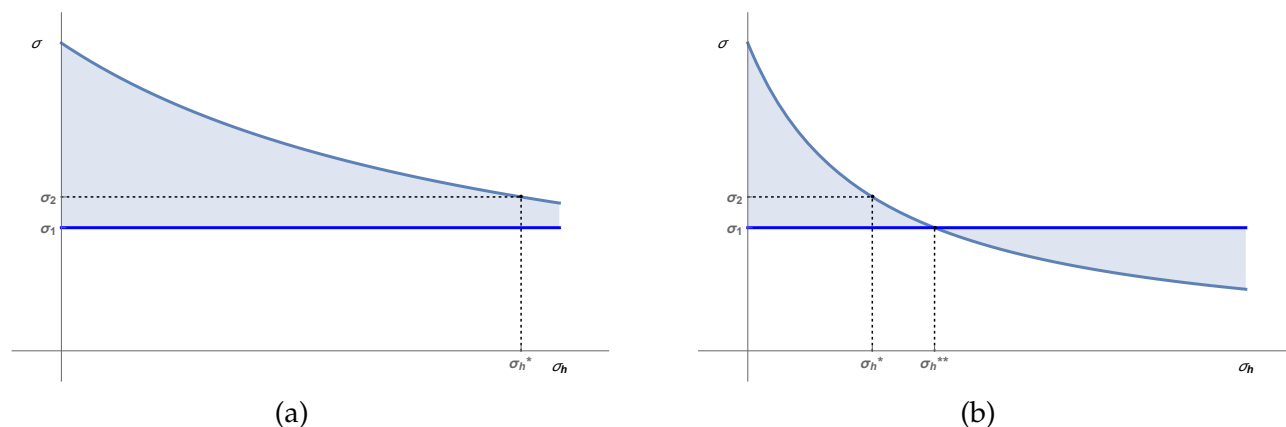
The above equation is satisfied as long as $(1 - \chi)R_s > R_i$, which is the assumption. Hence the coexistence equilibrium is always preferred to the traditional banking-only equilibrium.

1.3.4 Comparative Statics

It is interesting to see how the equilibrium depends on the fraction of banks with low innovation costs, which captures the degree of popularization of the financial technology. When the mass of innovative banks is small, they will be crowded out due to a prohibitively high interest rate in a market plagued by excessive risk taking. As the mass of innovative banks increases, there will be a separating equilibrium in which only innovative banks participate in banking. However, it is eventually replaced by a pooling equilibrium, in which non-innovative banks enter the market to exploit the low interest rate. Note that the deterioration of asset quality occurs only when the fraction of inno-

vative banks is sufficiently large that non-innovative banks can “ride the boom”. This prediction is consistent with the anecdotal evidence/pattern that the deterioration of asset quality coincides with relatively high average quality.

Figure 1.2: Comparative Statics: θ



1.4 Shadow Bank Regulation

The aim of the shadow bank regulation is to reduce systemic risk and thus provide a commitment tool to not provide a liquidity backstop to shadow banks. I show that the combination of inspection and punishment prevents non-innovative banks from taking excessive risks and achieves the first-best allocation even when the regulator has imperfect information.

1.4.1 Deterrent Effect of Designation

By requesting non-public information from banks, the regulator is capable of learning the underlying risks of banks’ asset holdings. By investing in its inspection and research capacity, it is able to identify the true types of a fraction $\eta \in (0, 1)$ of banks holding risky assets in period 1. The shadow bank regulator chooses (1) the capability of inspection η and (2) the severity of punishment Δ for identified i -banks. Under the regulatory framework (η, Δ) , each shadow banker has a probability η of being identified.

Examples of punishment measures include fines, forced liquidation/divestiture, and additional capital charges. For instance, since its designation, GE Capital has decreased its total assets by over 50 percent, shifted away from short-term funding, and reduced its interconnectedness with large financial institutions. Further, the company no longer owns any U.S. depository institutions and does not provide financing to consumers or small business customers in the United States. Another example is AIG, which has reduced the amounts of its total debt outstanding, short-term debt, derivatives, securities lending, repurchase agreements, and total assets. The company has sold certain non-core businesses, such as its aircraft leasing and mortgage guaranty businesses, and reduced its risk.

Proposition 1 (Optimal Designation Policy). *Given the inspection capacity η , the deterrence effect of shadow bank regulation is effective as long as the additional punishment is larger than the threshold value $\Delta^*(\eta)$, where*

$$\Delta^*(\eta) = \frac{1-\eta}{\eta}(1-q)(R_i - r_{SB}) - \frac{1}{\eta}(R_f - r_{TB})$$

Proof. Given the regulatory framework, an i -bank will be identified in period 1 with probability η . However, if undetected, it will secure a banking license to operate in the shadow banking sector. A non-innovative bank prefers traditional banking to shadow banking only when the profits of the former are larger than the expected profits of regulatory arbitrage.

Hence an effective inspection capacity η satisfies:

$$(1-\eta)(1-q)(R_i - r_{SB}) + \eta(0 - \Delta) < R_f - r_{TB} \quad (1.4.1)$$

The punishment in the shadow banking sector should be effective in ensuring that only superior risky assets will be financed using s-claims. Given the interest rate r_{SB} ,

banks will be deterred from excessive risk-taking if $\Delta > \frac{1-\eta}{\eta}(1-q)(R_i - r_{SB}) - \frac{1}{\eta}(R_f - r_{TB}) \equiv \Delta^*$.

□

A regulation is defined as optimal if the regulated market achieves the first-best allocation. The optimal regulation is not limited to the implementation of a direct revelation mechanism, where banks tell the truth about their asset types. As long as the low type (here, i -banks) are prevented from pooling with the high type, the regulation is optimal.

Proposition 2 (Optimal Shadow Bank Regulation). *The shadow bank regulation with effective deterrence prevents non-innovative banks from taking excessive risks and achieves the first-best allocation.*

Low-cost banks have no incentives to deviate from the equilibrium since the profit from deviation is lower. High-cost banks would like to deviate by participating in shadow banking but are unable to do so under regulation. Hence, low-cost banks will indeed innovate, financing superior risky projects and raising funding in the shadow banking sector. High-cost banks do not innovate and operate in the regulated traditional banking sector.

The insight is that shadow banking, defined as a sector issuing risky financial claims, is part and parcel of the optimal regulatory design to foster innovation. Since private and social interests are aligned when innovation is undertaken, unregulated banking with entry approval is welfare improving.

For any supervision intensity $\eta \in [0, 1]$, there is a punishment level Δ^* that will deliver effective regulatory policy. The punishment power of the regulator over designated banks generates a deterrent effect and improves welfare when the inspection capacity is low. The supervision and designation can improve social welfare through the deterrent effect, the effectiveness of which depends on the regulator's inspection capacity η and the magnitude of the punishment Δ . The effective punishment can be larger than the negative externalities generated by the bad banks to deter excess risk-taking activities when

the regulator only has imperfect information and limited resources.

To effectively deter banks' excessive risk-taking, the lower the inspection capacity is, the larger the punishment should be. If the additional punishment's upper bound is binding, more resources should be spent on increasing monitoring and inspection capacity

Under the optimal policy, since banks are deterred from investing in inferior risky assets, the number of designated SIFIs is zero. Political pressure arises when inspection is government funded. Punishment has deterrence power only when combined with continuing monitoring effort, which is costly. Optimal designation policy incurs large costs with no fine revenue and thus becomes a convenient target for political criticism and abolition proposals.

The functional independence comes from fiscal independence. The shadow bank regulator should be able to independently determine its budget and staffing. To implement the optimal designation policy, the assessment fee should be imposed on nonbank financial institutions under review, regardless of whether they are designated.

Moreover, the supervision and punishment regulation can be applied to all types of financial intermediaries, so that the first-best allocation can be achieved in the complex shadow banking sector. Therefore, to make the deterrent effect work, the shadow banks must have sufficient capital or company assets to ensure credible punishments. For instance, transfers from parent companies, capital requirements, licenses and permits ensure that that a company cannot enter bankruptcy while leaving nothing left to be punished.

The rationale for a capital requirement is as follows: (1) It is more resilient to return volatility (loss-absorbing capital buffer); (2) it aligns public and private interests (prevents excessive risk-taking); and (3) it ensures credible punishment (under asymmetric information).

Heavy punishment is necessary to deter bad shadow banks by removing the opportunistic mindset and creating a credible threat. The effectiveness of deterrence depends

on the regulator's inspection capacity and punishment severity and may achieve the first-best allocation. Policy implications are discussed below regarding the role played by the FSOC and SIFI designation. This paper also analyzes the effect of investor composition on the design of financial contracts and regulation.

While regulatory arbitrage is believed to be the main driver of shadow banking activities, shadow bank regulation is more than closing regulatory gaps and differences. In contrast to traditional banks, the shadow banking system is complex and intertwined, spanning across all financial industries, including securities, insurance, asset management and various financial markets.

The liquidity shortage caused by maturity mismatch in financial intermediation can be mitigated by the central bank through public liquidity creation. Thus, a timely intervention by the central bank during a liquidity crisis can prevent financial contagion and pecuniary externalities due to asset fire sales by constrained financial intermediaries. Thus, why does the government not extend deposit insurance to all demandable financial claims that could suffer from a run? In fact, the U.S. government provided guarantees for the MMF industry shortly after the failure of Lehman Brothers in 2008.

1.5 Systemic Risk and Credibility of Designation

In this section I examine the case where the designation is sub-optimal: the punishment severity is smaller than the optimal case so that non-innovative banks still have incentives to operate in the shadow banking sector.

With the random inspection and imperfect designation punishment, banks choose to innovate if, $(1 - q)[(1 - \chi)R_s - r] > (1 - \eta)(1 - q)(R_i - r) + \eta(0 - \Delta)$, namely when

$$\chi < 1 - \frac{R_i + \frac{q}{1-q}R_i - \eta(R_i - r_{SB}) - \frac{1}{1-q}\eta\Delta}{R_s} \equiv \hat{\chi}' \quad (1.5.1)$$

Denote θ' the fraction of innovative banks under the new innovation strategy. Now

it is a function of not only the payoffs of risky assets, but also the designation policy, the punishment severity, and the interest rate in the shadow banking sector.

In the coexistence equilibrium, the size of shadow banking is $\sigma'_{SB} = \theta' + (1 - \theta')\sigma_h$, while the size of traditional banking is $\sigma'_{TB} = (1 - \theta')(1 - \sigma_h)$.

Again, the interest rate in the coexistence equilibrium thus makes high-cost banks indifferent between participating in traditional or shadow banking:

$$r'_{SB} = \frac{1}{1 - \eta} \frac{1}{1 - q} r_{TB} - \left[\frac{1}{1 - \eta} \frac{R_f + \eta\Delta}{1 - q} - R_i \right] \quad (1.5.2)$$

According to households' participation constraint, the interest rate should satisfy:

$$r'_{SB} = \frac{1}{(1 - q) \left[1 - \left(1 - \frac{\theta'}{\theta' + (1 - \theta')\sigma'_h} \right) \eta \right]} \left[r_{TB} - q \frac{\theta'}{\theta' + (1 - \theta')\sigma'_h} R_a - \left(1 - \frac{\theta'}{\theta' + (1 - \theta')\sigma'_h} \right) \eta L \right] \quad (1.5.3)$$

Combining equation 1.5.2 and 1.5.3, we obtain the equilibrium σ_h^{**} :

$$\sigma_h^{**} = \frac{\theta'}{1 - \theta'} \left[\frac{(1 - \eta)q(1 - \chi)R_a}{R_f - (1 - q)R_i + \eta[(1 - q)R_i + \Delta - r_{TB}]} - 1 \right] \quad (1.5.4)$$

1.5.1 Maturity Mismatch and Liquidity Risk

To generate strategic complementarity in a model in which one bank is matched with one household, I impose a sequential selling constraint in the asset market in period 1, where banks unable to borrow from sophisticated investors line up to sell their assets.

Denote $\varphi = \int_i \varphi_i$ as the total amount of liquidation of risky assets in period 1. I assume that when $\varphi \geq \tilde{\varphi}$, where $\tilde{\varphi} \in ((1 - \theta)\lambda, \lambda)$, the crisis state will be triggered, where the payoffs of risky assets will be postponed to a remote period 3. In the interbank lending market, if sophisticated investors observe $\varphi < \tilde{\varphi}$, they will lend to banks holding risky assets. When $\varphi \geq \tilde{\varphi}$, adding the flavor of pecuniary externalities and a mark-to-market borrowing constraint, I assume a *sequential lending constraint*: Only the first $\theta\tilde{\varphi}$

banks will be able to close the deal, while the remaining banks in line fail to find any lenders accepting risky assets as collateral. The frozen interbank market forces the banks to file for bankruptcy, leading to more asset liquidation.

I solve for the subgame perfect equilibrium through backward induction: First, given households' beliefs and redemption strategy, derive a bank's innovation and asset holding strategies. Then, given the mix of assets in the banking sector, determine households' redemption strategy and assess whether the households' participation constraint is satisfied. To close the solution to the equilibrium, calculate banks' expected profits to assess banks' participation constraints.

Given their beliefs, patient households decide whether to demand early redemption. When the belief is $\varphi < \tilde{\varphi}$, a patient household who chooses to wait will obtain a higher expected payoff in period 2 since $r_{LF} \geq 1$. However, if the belief is $\varphi \geq \tilde{\varphi}$, patient households will demand early redemption: They will receive one unit of funding if their banks are the first $\tilde{\varphi}$ in line to borrow in the interbank market and zero if their banks are not among the first $\tilde{\varphi}$ banks or if they choose to wait until period 2. Therefore, patient households will adopt a threshold strategy: Demand early redemption if $\varphi > \tilde{\varphi}$, and wait if $\varphi < \tilde{\varphi}$.

The liquidation of inferior risky assets has negative externalities. Upon liquidation, the underlying projects are discontinued, which has an adverse effect on the productivity of projects composing superior risky assets and as a result reduces the market price of superior risky assets. For instance, the fire sale of a house due to the low-income owner's delinquency on his mortgage will lower the prices of all other homes on the block, even when these owners have sufficient incomes to meet the mortgage payment. However, when the market price of houses drops below the remaining mortgage payments ("under water"), good owners may strategically default on their payments, leading to a new round of delinquencies and fire sales and eventually a crisis in which superior risky assets are worthless.

The negative feedback loop induces strategic complementarity among households and leads to a self-fulfilling bank run, where the interbank market collapses, market liquidity evaporates, all assets are liquidated, and the economy is in a crisis state. Despite the different setting, my model derives the same sunspot equilibria as in [Diamond and Dybvig \(1983\)](#), where the self-fulfilling prophecy results from the sequential service constraint of a bank matched with a continuum of households. Here, the strategic complementarity results from the negative externalities in asset liquidation and contributes to the multiplicity of equilibria. The good equilibrium can be sustained when only impatient households demand early redemption and banks holding superior risky assets are able to borrow from the interbank market. The total amount of liquidation is $(1 - \theta)\lambda$, and the crisis state is not triggered. However, there is also a bad equilibrium of self-fulfilling runs, where all households demand early redemption and the interbank market freezes, which forces banks to liquidate assets and triggers the crisis state.

The self-fulfilling crisis captures the liquidity evaporation in the interbank lending market, when lenders become extremely cautious and conservative in purchasing such financial assets or providing funding to such types of banks. The contagion effect among banks and the asset liquidation channel of systemic risk are salient features of the 2007-2009 financial crisis and the emphasis of post-crisis regulatory evaluation.

When low-cost banks innovate, the liquidity backstop breaks even in expectation if $\sigma(1 - \chi_l)[(1 - q)R_s + qR_a] + (1 - \sigma)(1 - q)R_i \geq 1$, which requires

$$\sigma \geq \frac{1 - (1 - q)R_i}{(1 - q)(1 - \chi_l)R_s + q(1 - \chi_l)R_a - (1 - q)R_i} \quad (1.5.5)$$

Note that the right hand side of the inequality is smaller than $\sigma_2 = \frac{[R_f - (1 - q)R_i]}{q(1 - \chi_l)R_a}$.

When low-cost banks do not innovate, the liquidity backstop never breaks even. However, regardless of whether the low-cost banks innovate, the regulator has to provide a liquidity backstop to prevent the crisis state, which presents a time-inconsistency

problem in public liquidity injection

While banks that choose to innovate prefer superior risky projects, non-innovative banks deviate from the first-best allocation by investing in inferior risky assets and taking excessive risks. The deviation makes the composition of assets in the financial industry riskier, which has two adverse effects on social welfare. First, it lowers the expected payoffs in the good state, since inferior risky assets yield less than safe assets in expectation. Second, it increases the probability of the bad state because the inferior risky assets contaminate the asset liquidation market and may trigger a liquidity seize-up.

Expecting the regulator to step in and provide guarantees on and extend loans to shadow banks, households then have no incentives to run on the banks. There will be no liquidation risks, but the bailout expectation will induce a moral hazard problem and mispricing in the asset market. Shadow banks are able to finance inferior risky assets without facing a penalty or market discipline. The accumulation of systemic risk in the shadow banking sector reinforces the bailout expectation.

I have derived the banks' strategies given the households' strategy and beliefs. To complete the equilibrium, I will derive the households' redemption strategy in this section. When $\sigma_l = \sigma_h = 0$, only the traditional banking sector exists. The public liquidity provision is riskless, and only impatient households will demand early redemption. When $\sigma_h > 0$, however, *s*-banks are pooled with *i*-banks in the shadow banking sector. When the regulator does not provide the liquidity backstop to shadow banks, shadow banks facing early redemption requests have to liquidate their assets.

In the bad equilibrium, without the liquidity backstop provided by the regulator, shadow banks have to liquidate their assets, which leads to fire sales and triggers the crisis state, where the payoffs of assets are postponed to a remote, indefinite period 3. A fraction $1 - \tilde{\varphi}$ of households, patient or impatient, cannot consume in period 1 or period 2, because the market freezes and the assets cannot be sold.

The welfare-maximizing regulator has to step in, taking unsold assets and extending

loans to these banks so that households will be able to consume. Given the inevitability of ex post rescue, the regulator could do better by extending its LOLR function to shadow banks prior to the triggering of crisis state.

With the liquidity backstop provided by the regulator, there will be no forced liquidation in period 1, and the crisis state will not be triggered. A household demanding early redemption will always receive r^1 , and a household that does not demand early redemption expects to receive $(1 - q)r + \sigma(1 - \chi_l)R_a$ in period 2.

The difficult task faced by the regulator results primarily from asymmetric information. While financial intermediaries know clearly whether they are conducting beneficial innovation or disguising themselves as innovative banks, the regulator, which has less information regarding the nature of financial activities, cannot immediately identify whether the the risky claims are backed by superior or inferior risky assets. For instance, suppose now that mortgages are extended to people with lower FICO scores, who, based on historical statistics, would have been rejected. This could be because the financial intermediaries have applied more advanced technology than FICO scores to assess the credit risks of borrowers (SOFI), but it could also be the case that the new form of lending is used to mask the poor quality of underlying assets.

The regulator faces a dilemma between fostering financial innovation and preventing excessive risks when considering extending the capital regulation to the whole financial industry¹². When innovation technology is not available (i.e., $\theta = 0$), the only risky assets available are inferior ones. The first-best allocation can be achieved by prohibiting risk-taking in the traditional banking sector and prohibiting all shadow banking activities. In equilibrium, banks raise funding through regulated traditional banking (the only banking sector), and only safe projects are funded.

¹²Technically, subjecting non-bank financial institutions to capital requirements involves many industry-specific difficulties. For example, MMFs are financed primarily through capital, which satisfies the capital requirement, but since the capital is mainly demandable equity, it has little loss-absorbing capacity and may even trigger a liquidity crisis when it breaks the buck, as exemplified by the Reserve Primary Fund in 2008.

When innovative technology is introduced (i.e., $\theta \in (0, 1]$), however, the aforementioned conventional regulation in the banking industry may not be optimal, since it prevents excessive risk-taking at the cost of forgoing beneficial innovation opportunities. Inferior risky assets can also disguise themselves under the cover of financial innovation. An incompetent regulator that fails to curb excessive risk-taking by non-innovative banks will also discourage innovative banks from investing in welfare-improving innovation technology.

1.6 The Financial Stability Oversight Council (FSOC)

During the 2007-2009 financial crisis, the financial distress at certain non-bank financial companies contributed to a broad seizing up of financial markets and stress at other financial firms. Many of these non-bank financial companies were not subject to the type of regulation and consolidated supervision applied to bank holding companies (BHCs), nor were there effective mechanisms in place to resolve the largest and most interconnected of these non-bank financial companies without causing further instability.

To address any potential risks to financial stability posed by these companies, the DFA established the FSOC, which brings together federal and state financial regulators to look across the financial system to identify risks to financial stability, promote market discipline, and respond to emerging threats to the stability of the U.S. financial system.

The designation authority distinguishes the FSOC from any other collaborative body of financial regulators, which fills the regulatory gap and improves policy coordination. Authorized by the DFA, the FSOC has the statutory power to designate any non-bank financial institutions posing a threat to the U.S. financial system as SIFIs. The designated SIFIs will be supervised by the Board of Governors of the Federal Reserve and subject to enhanced prudential standards.

1.6.1 SIFI Designation

A designation means that the FSOC has determined that “the company’s material financial distress, or the nature, scope, size, scale, concentration, interconnectedness, or mix of the activities of the company could pose a threat to U.S. financial stability.”

Since the establishment of the FSOC, four companies have been labeled SIFIs: American International Group (AIG) on July 8, 2013, General Electric Capital Corporation (GECC) on July 8, 2013, Prudential Financial on September 19, 2013, and MetLife on December 18, 2014. These non-bank financial institutions are large, complex, interconnected with other major financial intermediaries and comparable to the largest U.S. BHCs in terms of size of assets and nature of business. For instance, AIG was the third-largest insurance company in the United States and one of the largest insurers in the world, operating across many different markets. Prior to the onset of the financial crisis, the company expanded its operations to include non-insurance businesses. While the company’s strategy, funding profile, and global footprint have changed greatly since the financial crisis, AIG remains a large and complex company with meaningful non-insurance-related exposures. During the intensification of the financial crisis in the fall of 2008, AIG became the recipient of considerable government support, which was deemed necessary to avoid an even larger financial disruption.

First, these non-bank financial institutions have the same maturity mismatch problem as traditional banks do and thus are subject to runs. This is either due to heavy reliance on the wholesale short-term funding markets or because of their offering of financial products with early withdrawal features. For example, GECC was a significant issuer of commercial paper (CP) in the United States and was subject to rollover risks; MetLife’s funding agreement-backed commercial paper (FABCP) constitutes a significant portion of the company’s capital market financing activities, which is short-term and exposes MetLife to investment renewal risks; a significant amount of Prudential’s U.S. life

insurance policies are subject to early withdrawal and include a significant cash surrender value; and many of AIG's life insurance and annuity products, while intended to be long-term liabilities, have features that could make them vulnerable to rapid and early withdrawals by policyholders.

Second, asset liquidation imposes significant externalities on other financial intermediaries and markets, particularly during a period of overall stress in the financial services industry and in a weak macroeconomic environment, when liquidity dries up and price swings can be magnified. For instance, material financial distress at GECC could trigger a run on MMFs and lead to a broader withdrawal of investments from the CP market and other short-term funding markets. The liquidation of a significant portion of Prudential's assets could cause significant disruptions to key markets including the corporate debt and asset-backed securities (ABS) markets, and the severity of the disruption caused by a forced liquidation of Prudential's assets could be amplified by the fact that the investment portfolios of many large insurance companies are composed of similar assets, which could cause significant reductions in asset values and losses for those firms. A rapid liquidation of AIG's life insurance and annuity liabilities could strain AIG's liquidity resources and compel the company to liquidate a substantial portion of its large portfolio of relatively illiquid corporate and foreign bonds, as well as ABS.

Third, the negative effects are amplified by the strategic complementarity among market participants who have lost confidence in the financial strength of companies with similar products or balance sheet profiles. For instance, in addition to the direct effects of asset liquidation on financial markets, wide-ranging and rapid withdrawals by AIG policyholders and the associated deterioration of AIG's financial condition could cause financial contagion if the negative sentiment and uncertainty associated with material distress at AIG spreads to other insurers. In particular, if distress at AIG were to cause concerns among policyholders at other insurers, those insurers could experience unanticipated increases in surrender activity that could strain liquidity resources, potentially

impairing the financial condition of multiple insurers across the industry, causing significant damage to the broader economy.

1.6.2 Enhanced Prudential Standards

A designation by the FSOC does not provide a non-bank financial company with any new access to government liquidity sources or create any authority for the government to bail out the company if it fails. To the contrary, the DFA seeks to eliminate taxpayer-funded bailouts of failing companies by subjecting them to stringent regulation prior to material failure.

For example, since its designation, GE Capital has fundamentally changed its business. Through a series of divestitures, a transformation of its funding model, and a corporate reorganization, the company has become a much less significant participant in financial markets and the economy. GE Capital has decreased its total assets by over 50 percent, shifted away from short-term funding, and reduced its interconnectedness with large financial institutions. Further, the company no longer owns any U.S. depository institutions and does not provide financing to consumers or small business customers in the United States.

Another example is AIG, which has reduced the amounts of its total debt outstanding, short-term debt, derivatives, securities lending, repurchase agreements, and total assets. Capital market exposures to AIG have decreased, and the company has sold certain businesses in which it held dominant market shares, rendering the company less interconnected with other financial institutions and smaller in scope and size. The company's focus on traditional insurance activities and its wind-down of non-core businesses, such as the aircraft leasing and mortgage guaranty businesses, have reduced its risk.

As a result, the FSOC voted to rescind the designation of GE Capital on June 28, 2016, that of AIG in 2017, and that of Prudential Financial on October 16, 2018¹³. Currently, the

¹³MetLife challenged its SIFI determination with a lawsuit filed in the US District Court for the District of

number of designated SIFIs is zero.

1.6.3 Identifying Emerging Risks

As history has shown, including in 2008, financial crises can be difficult to predict and have consequences that are both far-reaching and unanticipated. Consistent with its mission to identify potential threats before they occur, the FSOC focuses on the potential consequences of material financial distress at non-bank financial institutions, especially the transmission of the negative effects of a non-bank financial company's material financial distress to the financial system.

There are three main transmission channels identified in the Interpretive Guidance of the FSOC: exposure, asset liquidation and critical function or service¹⁴. To support the decision-making of the FSOC, the Office of Financial Research (OFR), created at the same time, is responsible for collecting data and conducting analysis to identify emerging risks in the financial system and can use subpoena power to fulfill its responsibilities when necessary. A significant amount of nonpublic information is collected and utilized to identify the risks posed by some financial activities, as well as to determine the designation of SIFIs.

The creation of the FSOC brought significant changes to the post-crisis regulatory regime. Not only has the regulatory perimeter, for the first time, been expanded to cover

Colombia in January 2015. The US District Court ruled in MetLife's favor in March 2016, which the FSOC appealed. However, after the change in the White House in January 2017, stays were issued in the case, including an order holding the appeal in abeyance until the issuance of a report required to be made by the Treasury Department, pursuant to an April 2017 Presidential Memorandum, after the Treasury Secretary undertook a thorough review of the FSOC determination process. The case was eventually dismissed in January 2018.

¹⁴Specifically, the FSOC evaluates whether (1) a non-bank financial company's creditors, counterparties, investors, or other market participants have exposure to the non-bank financial company that is significant enough to materially impair those creditors, counterparties, investors, or other market participants and thereby pose a threat to U.S. financial stability; (2) a non-bank financial company holds assets that, if liquidated quickly, would cause a fall in asset prices and thereby significantly disrupt trading or funding in key markets or cause significant losses or funding problems for other firms with similar holdings; and (3) a nonbank financial company is no longer able or willing to provide a critical function or service that is relied upon by market participants and for which there are no ready substitutes. For the purpose of my model, I focus on the asset liquidation channel of systemic risks posed by these financial institutions.

all financial institutions including shadow banks, but the implementation approach has been shifted from industry- to entity-based regulation. The macro-prudential regulator is substantially empowered and is expected to reduce government bailout expectations and help prevent another financial crisis.

Recognizing the FSOC's potential as a macro-prudential regulator becomes even more valuable when there are political trends in rolling back regulation. In recent years, arguments against the FSOC (and the DFA at large) are increasing under four main points. First, the expansion of the regulatory umbrella will impede financial innovation and hence make it more difficult for firms to obtain funding and lead to a slower recovery and a less vigorous economy. Second, the FSOC does not reduce, and will even reinforce, the market's expectation of government bailouts due to SIFI designation. Third, burdensome regulation induces financial companies to reduce the variety of financial products and services they offer and will disproportionately harm small banks and customers. Finally, FSOC designation weakens the global competitiveness of large financial companies by placing heavier regulatory burdens on them.

I address these comments in my model and provide novel arguments for the recent policy debates. I note that regulation of shadow banking is actually a necessary condition to foster good financial innovation, which would otherwise be crowded out by bad actors under asymmetric information. Moreover, by focusing on shadow banks with large systemic risks and taking preemptive action before they become "too systemic to fail", the FSOC credibly reduces probability of future bailouts using taxpayers' money while relieving small- and medium-sized financial institutions of compliance costs. Furthermore, the regulation protects unsophisticated households, which may not be able to distinguish between good risky and bad risky assets and may be exploited as a result. By promoting beneficial financial innovation and reducing systemic risk, FSOC regulation can increase the resilience of the financial system and contribute to a healthier and more competitive economy in the long run, from which every market participant will benefit.

The prudential regulator is no longer in a reactive position; instead, it can take proactive action to gather information and identify underlying risks, substantially reducing expectations of a government bailout by containing systemic risk well before it poses an eminent threat to the whole financial system.

Empowering the regulator is endorsing financial innovation and stability. Lax regulation allows excessive risk-taking to crowd out beneficial innovation, eventually making unsophisticated households suffer and leading to costly government bailouts. If we recognize the tendency of the leveraged financial system to take excess risks and the severity of the too-systemic-to-fail problem, we will not be content with a regulator that is insensitive to the ever-evolving financial industry and remains uninformed. The regulator charged with macro-prudential responsibilities should be empowered and should not be apologetic about wielding regulatory power in service of the welfare of the whole economy.

1.7 Conclusions and Future Research

The conventional wisdom of tightening the regulation to curb excessive risk taking has been challenged since the financial crisis. Loosely regulated, the shadow banking system enables financial intermediaries to bypass regulatory requirements and weakens the effectiveness of financial regulation.

I propose a model where shadow banks are not necessarily bad for the economy and stress on the positive force of shadow banks in fostering innovation. The model captures the bright side of shadow banking and sheds light on the optimal regulation design. Due to the replacement effect *a la* Arrow, traditional banks hesitate to offer new financial services, leading to inefficiency in investment. Shadow banks, however, enter into the game with tech and financial innovations, which provides liquidity to firms in need. As the shadow banking sector grows, the benefits of financial innovation are gradually outweighed by the threats of systemic risk. Once regulated, they will not be able to innovate

as much due to costly capital charges. While the traditional banking sector is regulated to credibly provide safe claims, a sector operating outside traditional regulation creates space for innovative financial intermediaries to take risks. However, under asymmetric information, the sector is contaminated by non-innovative banks seeking to circumvent traditional regulation and take excessive risks, thus becoming a shadow banking sector.

The paper has important policy implications. First, the emergence of shadow banks does not necessarily call for immediate regulation tightening. The regulatory authority should recognize the usefulness of the unregulated sector and rethink the boundary of regulation. Second, my paper is the first to formalize the designation authority of the Financial Stability Oversight Council over non-bank financial institutions in controlling risk accumulation in the shadow banking sector. The regulator, with imperfect information and limited resources, faces a tradeoff between tightening regulation to maintain stability and loosening regulation to foster innovation. Third, I propose a new approach to shadow bank regulation, highlighting the deterrence effect of inspection and punishment. Levying heavy punishment on shadow banks detected as taking excessive risks is critical to deliver effective regulatory deterrence, which justifies burdensome regulation on the designated systemically important financial institutions. The punishment has macroprudential impacts not only by controlling the existing systemic risks of the designated but also by deterring excessive risk-taking behavior *ex ante*. Future research directions include endogenizing the fire sale pricing in the asset market as well as empirical studies on the deterrent effect.

Chapter 2

Bank Liability Structure and Optimal Deposit Insurance

2.1 Introduction

Traditional wisdom puts deposit insurance on the opposite side of market discipline, arguing that the introduction of deposit insurance leads to moral hazard problem and will induce more risk-taking by banks. This argument guides the discussion on the role of deposit insurance and the design of risk-adjusted insurance premium for the purpose of mitigating the negative effects.

However, should we take the market discipline prior to deposit insurance for granted? My answer is no. By introducing agents with heterogeneous ability to acquire bank asset information, I show that the deposit insurance on the uninformed funding can benefit the market discipline efforts exerted by the informed agents. This provides a new rationale for capped deposit insurance and the optimal design of the insurance coverage limit, which goes beyond the cost-benefit analysis between preventing bank runs and reducing payout costs.

In my model, agents only differ in the ability of obtaining bank asset information. In the exogenous case, the quality of bank assets is revealed to only a fraction of the agents, who can then decide whether to act on that piece of information. In the endogenous information acquisition scenario, agents who differ in wealth choose whether to exert monitoring efforts and acquire more precise information regarding the bank's balance sheets. However, given the fixed monitoring cost, only wealthy depositors will choose to exert efforts and become informed with regard to the underlying assets.

The leveraged bank has incentives to take excess risks by choosing bad assets, which are riskier and less socially desirable than good ones. The market discipline, defined as the agents' efforts to curb the bank's excessive risk-taking incentives, can be implemented by a threat-to-run strategy by the informed agents. The strategy depends on the quality of the bank assets revealed to informed agents in the interim period. The informed will demand early withdrawal when the underlying assets are of bad quality and will wait

when the assets are good. When the underlying assets are of good quality, the informed agents choose to wait (and even lend more money) to stabilize the bank should there be a panic-based run among uninformed depositors. The stabilizing effect of informed funding creates an incentive for banks to discipline their investment decisions and hold good assets, counting on the informed depositors come to their rescue during runs.

Due to the strategic complementarity in bank runs, the benefits of the information advantage are affected by the bank's liability structure, which in this case refers to the relative size of informed and uninformed funding. When the size of liquidity that can be deployed by the informed agents is relatively small, a destabilizing force emerges: even when the underlying assets are good, the liquidity demand from the uninformed investors are so large than the bank has to liquidate (some of) the assets. This undermines the return prospects of informed agents, causing them to forgo superior information and run on the bank as well. Expecting this, depositors may choose not to acquire any additional information regarding the assets of the bank. The discouragement from exerting monitoring efforts weakens the market discipline and leads to efficiency loss.

The model has two insights: First, it shows that market discipline efforts do not naturally take place, even when some of the agents are handed with the necessary information. Without deposit insurance, there can be no market discipline from the agents anyways. Second, it implies that a capped deposit insurance, by stabilizing relatively small- and medium-sized depositors of the bank, can restore monitoring incentives of the large-sized depositors and benefit market discipline. Given the complementarity between strategies of informed and uninformed agents, eliminating bank runs initiated by uninformed agents restores incentives for other agents to obtain private information and use the short-term contract as a discipline tool.

My paper contributes to the literature on banking and applied corporate finance theory besides offering a new angle in terms of optimal design of deposit insurance.

My paper builds on [Goldstein and Pauzner \(2005\)](#)'s techniques in obtaining a unique

threshold equilibrium, with the emphasis on the market discipline and the monitoring efforts through information acquisition, which was not studied in previous literature. Their paper derives a unique equilibrium to pin down the probability of bank runs. Their focus is on the optimal design of demand deposit contract with the tradeoff between the benefits of risk sharing and the costs of more bank runs.

There are three distinctions in my model: The return is jointly determined by the state of the nature (economic fundamentals), θ , and the type of the loans, η . This distinguishes my paper from Goldstein and Pauzner (2005), where the return R is a constant and the state of the nature affects expected payoff only through the probability of success $p(\theta)$. Another difference is that in my model, the bank chooses to maximize its profit, instead of the assumption in Goldstein and Pauzner (2005) (and since Diamond and Dybvig (1983)) that banks are mutually owned by depositors and retain no residual value at the end of the period. The changes in model setting enables the analysis on moral hazard and market discipline, which is at the center of the debate on the optimal design of deposit insurance.

Deposit insurance has been a classic topic in economic research since the Great Depression and has attracted attention from scholars and practitioners. There are three main strands of literature: The first strand addresses the nature of deposit insurance, viewing it as a put option and thus applying the state-of-art option pricing theory (Stoll (1969), Black and Scholes (1973), Merton (1973a), Merton (1973b)) to evaluate the cost of the insurance. For instance, Merton (1977) demonstrates an isomorphic correspondence between loan guarantees and common stock put options and then uses the well-developed theory of option pricing to derive the formula. Several other papers deal with more specific pricing mechanism about the risk-adjusted deposit insurance (Ronn and Verma, 1986).

The second strand examines market settings where deposit insurance can (or cannot) be welfare-improving. Diamond and Dybvig's influential paper in 1983 shows circumstances when government intervention through providing deposit insurance can produce

superior contracts and therefore increase social welfare. They admit, however, that the riskless technology they used in the model abstracts from potential moral hazard problem since there is no room for bank managers to select the risk of bank portfolios in an unobserved way, and that “introducing risky assets and moral hazard would be an interesting extension of our model”. For instance, it has been analysed in complete market settings where deposit insurance is redundant and can provide no social improvement (Kareken and Wallace (1978), Dothan and Williams (1980)).

The third, and perhaps the most important strand of literature, focuses on the impacts of deposit insurance on financial stability, using both theoretical and empirical approaches. Deposit insurance, on one hand, makes the deposit a real riskless asset and enhances depositors’ confidence, therefore eliminating the bad equilibrium of bank runs. On the other hand, like other insurances, due to asymmetric information (here, partly unobserved bank managers risk-taking behaviors), the flat-rate deposit insurance gives the bank managers incentives to pursue riskier investments with higher-expected returns, hence making the financial system exposed to more risks. Demirgüç-Kunt and Kane (2002) found evidence from cross-country data emphasizing on the importance of assessing and remedying “weaknesses in their informational and supervisory environments” before a country adopts explicit deposit insurance. Anginer, Demirgüç-Kunt and Zhu (2014) analyzed the comparative magnitude of the two effects of deposit insurance in different times. One of their conclusions is that the moral hazard effect of deposit insurance dominates in good times while the stabilization effect of deposit insurance dominates in turbulent times. Additionally, they found that good bank supervision can alleviate the unintended consequences of deposit insurance on bank systemic risk during good times. Others, Keeley (1990) tests the hypothesis that increases in competition caused bank charter values to decline, which in turn caused banks to increase default risk through increases in asset risk and reductions in capital.

2.2 The Baseline Model

2.2.1 The Setup

In the three-period ($t = 0, 1, 2$) model, there are a continuum of agents¹, each endowed with one unit of wealth in period 0. Agents face preference shocks as in Diamond-Dybvig: their preference types will only be revealed in period 1. A fraction λ of the depositors is impatient, who only value early consumption c_1 , while the rest $1 - \lambda$ depositors are patient and indifferent between early and late consumption.

The underlying projects take two periods to finish and can be either "good ($k = G$)" or "bad ($k = B$)", which cannot be observed by depositors *ex ante*. If the project is interrupted in period 1, the liquidation value is assumed to be one unit of initial investment².

The long-run return is jointly determined by the state of the nature (economic fundamentals), θ , and the type of the loans, k . The state of the nature θ is drawn at the beginning of period 0 but is not revealed publicly. Once finished, type- k projects return R^k per dollar invested with probability $q^k(\theta)$ and zero otherwise. I assume θ follows a uniform distribution over the interval $[0, 1]$.

The expected payoff $q^k(\theta)R^k$ depends on both the types of loans and the state of the nature. Specifically, I assume that good assets have higher expected payoffs than bad ones, regardless of the realization of the state of nature:

¹ Here, the households represent depositors with account balance below the deposit insurance limit. Given account splitting and brokered deposit are not uncommon in reality, it is not that unrealistic to assume full protection for depositors. Moreover, the design of the model leaves adequate room for analyzing depositors who are not fully insured since they can be classified into the uninsured creditor group in the model. Therefore, by assuming that depositors are fully insured, the model is greatly simplified yet doesn't compromise its ability to shed light on empirical questions.

² I allow for variation in liquidation value in the extension, where the liquidation value of the project is $(1 - \delta)A$, where δ characterizes the cost of asset liquidation. Alternatively, we can think of δ as the haircut rate when the bank uses the project as collateral to borrow money and the collateral is evaluated at its book value. For now, δ is a constant to simplify analysis. In the context of fire sale, as what happened during the Financial Crisis of 2007-2009, it would be more appropriate to model the liquidation cost as an increasing function of the amount of assets liquidated, probably a non-linear one.

Assumption 4. Given the state of the nature θ , (1) Good projects have higher expected payoffs than bad projects: $q^G(\theta)R^G > q^B(\theta)R^B$; (2) The probability of success of good projects is higher than that of bad projects: $q^G(\theta) > q^B(\theta)$;

First-best allocation. The social planner knows the preference types of agents and quality of assets. It maximizes the social welfare, defined as the sum of utilities:

$$\max_{c_1^m, c_2^m, c_1^p, c_2^p} SW = \mathbb{E}[\lambda u(c_1^m + c_2^m) + (1 - \lambda)u(c_1^p + c_2^p)] \quad (2.2.1)$$

subject to the resource constraint

$$\lambda c_2^m + (1 - \lambda)c_2^p \leq \begin{cases} [1 - \lambda c_1^m - (1 - \lambda)c_1^p]R^\eta & \text{with prob. } q_k(\theta) \\ 0 & \text{with prob. } 1 - q_k(\theta) \end{cases} \quad (2.2.2)$$

In the first-best allocation, $c_2^{m*} = c_1^{p*} = 0$, $u'(c_1^{m*}) = q_\eta(\theta)R^\eta u'(c_2^{p*})$, and $\eta^* = G$. The consumption of patient agents in period 2 is $c_2^{p*} = \frac{1 - \lambda c_1^{m*}}{1 - \lambda} R^G$ with probability $q_G(\theta)$, and zero otherwise.

2.2.2 Demand Deposits and Bank Runs

From now on, I introduce a representative bank which raises funding from agents through demand-deposit contracts and make loans to firms. The demand deposit contract (r_1, r_2) promises to pay r_1 per unit of wealth deposited in period 0 to depositors who withdraw in period 1 (early withdrawal) and r_2 to those who wait until period 2. In order to attract depositors, banks will set $r_2 \geq r_1 \geq 1$.

Denote n the fraction of depositors demanding early withdrawal. The upper bound of r_2 is given by $\frac{(1 - nr_1)R}{1 - n}$. When $n < 1/r_1$, the bank is still able to stay in business, and the expected profit is $\pi(\theta, \eta)_{r_1, r_2, \eta} = q^\eta(\theta)[(1 - nr_1)R^\eta - (1 - n)r_2]$. The actual payment in period 2 \tilde{r}_2 is given by $\min\{r_2, \frac{1 - nr_1}{1 - n} R^\eta\}$.

Profit-maximizing banks may be tempted to take on bad projects due to the risk-shifting nature of debt contracts. That is, the bank cannot commit to making good loans. This motivates the use of short-term funding contract as a commitment device. Specifically, I assume $R^G < R^B$ and make the following assumption:

Assumption 5 (Excessive risk-taking incentives). *Given (r_1, r_2, θ) , the bank prefers to make riskier bad loans $q^B(\theta)[(1 - nr_1)R^B - (1 - n)r_2] > q^G(\theta)[(1 - nr_1)R^G - (1 - n)r_2]$*

My paper distinguishes from others in the following three aspects: First, I allow for socially undesirable assets to address the excessive risk-taking problem, instead of only one category of assets which assumes away the bank's investment choices.

Second, I change the demand-deposit contract so that the payment in period 2 is a pre-determined amount rather than the residual value of investment. The demand-deposit contract allows depositors to withdraw their money in both period 1 and period 2, with the one-period return of r_1 and two-period return of r_2 ³.

Third, the bank maximizes its own expected profit rather than social welfare. The new feature of the demand-deposit contract in my model places a wedge between the total payoff of projects and the return of long-term deposits. The bank chooses the interest rates specified in the demand-deposit contract, as well as the type of assets, to maximize profits.

Information structure. Agents and banks have asymmetric information regarding the fundamentals and the quality of bank assets. Different from agents, the bank is able to observe the fundamentals and distinguish project types at the beginning of period 0 and makes decisions accordingly.

Timeline. At the beginning of period 0, nature draws the fundamental θ from a uniform distribution over $[\underline{\theta}, \bar{\theta}]$. The bank observes the realized fundamentals immediately

³ In the appendix I examine the possibility of the bank offering two separate contracts to common and wealth depositors, namely different rates for retail and wholesale funding. The incentives of using short-term contracts still apply.

and chooses asset type k to maximize its expected profit. At the beginning of period 1, each agent receives a noisy private signal $\theta_i = \theta + \epsilon_i$, where ϵ_i is uniformly distributed over $[-\epsilon, \epsilon]$. In period 1, the preference shock is also realized. Agents learn their own preference types and decide whether to demand early withdrawal. The bank liquidates (some) assets to meet the liquidity demand. In period 2, the return is realized. The bank fulfills its payment obligations and retains the residual value (if any).

2.2.3 Benchmark: Equilibrium without Bank-asset Information

Consider the extreme case where all agents remain uninformed regarding the bank's assets. This brings us back to the [Goldstein and Pauzner \(2005\)](#) model, where a unique equilibrium is pinned down.

Given excessive risk taking incentives, the bank will always choose bad assets, namely $\eta = B$.

$$v(\theta, \{r_1, r_2\}) = \begin{cases} q_B(\theta)u(r_2) - u(r_1) & \text{if } n < \frac{1}{r_1} \\ 0 - \frac{1}{nr_1}u(r_1) & \text{if } n > \frac{1}{r_1} \end{cases} \quad (2.2.3)$$

When fundamental is good enough, banking is viable even when bad assets are invested. No incentives to monitor since the participation constraint of the agents is always satisfied.

2.3 Equilibrium under Exogenous Information Acquisition

Now suppose the bank asset information is exogenously revealed to a fraction ω of patient depositors in period 1, who are thus referred to as informed depositors. I assume that the bank asset information is perfect and reveals the true type of assets held by the bank. The strategy of informed depositors is dependent on both the bank asset type k and the state of the economy θ . These informed depositors are viewed as a collective which acts together, thus the fraction $(1 - \lambda)\omega$ can also be interpreted as the relative wealth held

by one representative patient depositor. The rest $(1 - \delta)(1 - \lambda)$ patient depositors are uninformed in terms of bank assets and only have a noisy signal about the fundamentals, θ_i .

The equilibrium consists of depositors' strategies of (1) withdrawal strategy: whether to demand early withdrawal,, $w_i = \{1, 0\}$, based on their preference types and private signals, and bank's strategies of (1) demand-deposit contract $D = (r_1, r_2)$, (2) portfolio strategy: which type of project to invest $\eta = \{B, G\}$, which satisfy the profit-maximizing, based on the signal θ .

Theorem 6. *When the fraction of informed agents ω is within a certain range, there is a unique equilibrium where informed agents run if the assets are bad ($\eta = B$) and do not run if the assets are good ($\eta = G$), regardless of the fundamentals, while uninformed agents run if they observe a signal below the threshold θ^* and do not run above. The bank choose good assets.*

The theorem can be used to calculate the proportion of early withdrawals. Denote A the event where the informed agents decide to run on the bank. According to the strategy, $\mathbb{1}_A = 0$ when $k = G$ and $\mathbb{1}_A = 1$ when $k = B$. The informed agents will not run in period 1 only if good assets are invested, while the uninformed agents who receive a signal below the threshold will run. Thus, the proportion of early withdrawals is determined by the fundamentals. Following Goldstein and Pauzner (2005), I use $n(\theta, \theta')$ to specify the proportion,

$$n(\theta, \theta') = \underbrace{\lambda}_{\text{impatient agents}} + \underbrace{(1 - \lambda)(1 - \omega) \text{prob}(\theta_i < \theta^*)}_{\text{uninformed}} + \underbrace{(1 - \lambda)\omega \mathbb{1}_A}_{\text{informed}} \quad (2.3.1)$$

patient agents

The signal θ_i follows a uniform distribution over $[\theta - \epsilon, \theta + \epsilon]$ since both the fundamentals θ and the noise ϵ are uniformly distributed. When $\theta < \theta^* - \epsilon$, the highest signal observed by agents is below the threshold, thus all uninformed agents will run on the bank, which means $n(\theta, \theta^*) = \lambda + (1 - \lambda)(1 - \omega)$. When $\theta > \theta^* + \epsilon$, even the

lowest signal observed by agents is above the threshold, therefore only impatient agents would demand early withdrawal, namely $n(\theta, \theta^*) = \lambda$. When θ is within the interval $[\theta^* - \epsilon, \theta^* + \epsilon]$, the proportion of early withdrawals is given by the linear function $n(\theta, \theta^*) = \lambda + (1 - \lambda)(1 - \omega)(\frac{1}{2} + \frac{\theta^* - \theta}{2\epsilon})$.

Given n , the bank will choose good assets only if $\pi_G > \pi_B$. I assume that the bank has incentives to take excessive risks when the informed agents choose to wait no matter which type of assets they observe, namely $\pi_B > \pi_G$ given $n|_{\mathbb{1}_A=0}$:

Assumption 7. When $\mathbb{1}_A = 0 \forall k$, $\pi_G - \pi_B < 0$, where

$$\pi_G = q_G[(1 - n|_{\mathbb{1}_A=0}r_1)R^G - (1 - n|_{\mathbb{1}_A=0})r_2]$$

$$\pi_B = q_B[(1 - n|_{\mathbb{1}_A=0}r_1)R^B - (1 - n|_{\mathbb{1}_A=0})r_2]$$

Denote $\Delta_{GB} \equiv q_G R^G - q_B R^B$ and $\Delta_q = q_G - q_B$. When $r_2 < \frac{\Delta_{GB}}{\Delta_q}$, the inequality holds for $n > \Delta_{GB} - r_2 \Delta_q r_1 \Delta_{GB} - r_2 \Delta_q$. When $r_2 > \frac{\Delta_{GB}}{\Delta_q}$, the inequality holds for all $n \in [\lambda, 1]$.

Definition 8 (Effective and credible market discipline). *The market discipline is effective if $\pi_G > \pi_B$ given the strategy of informed agents and $\pi_G < \pi_B$ without. The market discipline is credible if the aforementioned strategy satisfies the participation and incentive-compatibility constraints of the informed agents.*

In the model, it requires $\pi_G|_{\mathbb{1}_A=0} > \pi_B|_{\mathbb{1}_A=1}$ for the bank and $v|_{k=G} > 0$ and $v|_{k=B} < 0$ for the informed agents.

Definition 9 (Pivotal mass of informed agents). *The mass of informed agents is pivotal if $n|_{\mathbb{1}_A=1} > \hat{n}$ for any level of $n|_{\mathbb{1}_A=0} < \hat{n}$, where \hat{n} represents the mass of early withdrawals which leaves bank zero profits.*

When ω is pivotal, the threat-to-run strategy is effective since $\pi_B = 0$ and $\pi_G > 0$. To find out the smallest ω to implement effective market discipline, I focus on the case

where the mass of informed agents is non-pivotal, namely when ω satisfies $n|_{\mathbb{1}_A=1} < \hat{n}$ for any level of $n|_{\mathbb{1}_A=0} < \hat{n}$.

When ω is non-pivotal, effective market discipline requires $q_G[(1 - n|_{\mathbb{1}_A=0}r_1)R^G - (1 - n|_{\mathbb{1}_A=0})r_2] > q_B[(1 - n|_{\mathbb{1}_A=1}r_1)R^B - (1 - n|_{\mathbb{1}_A=1})r_2]$. Denote $\Delta = r_2(q_G - q_B) - r_1(q_G R^G - q_B R^B)$. The market discipline strategy is effective if the fraction of informed agents is larger enough, $\omega > \omega^*$, where

$$\omega^* \equiv \frac{[r_2\Delta q - \Delta_{GB}] - \Delta[\lambda + (1 - \lambda)\text{prob}(\theta_i < \theta^*)]}{(1 - \lambda)[(q_B R^B r_1 - r_2) - \Delta\text{prob}(\theta_i < \theta^*)]} \quad (2.3.2)$$

Now let's check whether the market discipline strategy is credible on the agents' side. Given n and η , patient agents, informed or uninformed, will choose to wait if the following differential is positive:

$$v(\theta, \eta, n, \tilde{n}) = \begin{cases} q_\eta(\theta)u(\tilde{r}_2) - u(r_1) & \text{if } n < \frac{1}{r_1} \\ 0 - \frac{1}{nr_1}u(r_1) & \text{if } n > \frac{1}{r_1} \end{cases} \quad (2.3.3)$$

where

$$\tilde{r}_2 = \min\left\{r_2, \frac{1 - nr_1}{1 - n}R^\eta\right\} = \begin{cases} r_2 & \text{if } \lambda \leq n \leq \hat{n} \\ \frac{1 - nr_1}{1 - n}R^\eta & \text{if } \hat{n} < n \leq \frac{1}{r_1} \\ 0 & \text{if } \frac{1}{r_1} < n \leq 1 \end{cases} \quad (2.3.4)$$

Three differences from the Goldstein and Pauzner paper: First, bank asset information η . This gives informational advantage to some agents in order to examine the market discipline mechanism when banks may take excessive risks. Second, the fraction of early withdrawals when informed agents run $\tilde{n} = n + (1 - \lambda)\delta$. The informed agents act collectively, and their withdrawal is no longer neglectible. Another interpretation is that the informed agents are wealthy people and have larger stakes in the bank, thus they are

motivated to acquire information or exert market discipline. Either way, the mass of informed agents matters to the stability of the bank. Third, the expected payment in period 2 \hat{r}_2 . It is the smaller value between the amount specified in the demand-deposit contract r_2 and the residual value $\frac{1-nr_1}{1-n}R^\eta$. This allows the bank to earn a profit; otherwise the banker's and the depositors' interests are always aligned, leaving no room for agency problem.

The effective market discipline requires that $v(\theta, G, n, \tilde{n}) > 0$ and $v(\theta, B, n, \tilde{n}) < 0$ for all θ . Note that $v(\theta, \eta, n, \tilde{n})$ is always negative when $\tilde{n} > \frac{1}{r_1}$. That is to say, informed agents will not run on good assets only under the prerequisite that $\tilde{n} < \frac{1}{r_1}$, which requires that the mass of informed agents is not too large.

$$\omega < \frac{\frac{1}{r_1} - n_0}{(1-\lambda)(1-\text{prob}(\theta_i < \theta))} \equiv \omega^* \quad (2.3.5)$$

where $n_0 = \lambda + (1-\lambda)\text{prob}(\theta_i < \theta^*)$, which is the extreme cases where all patient agents are uninformed about the bank asset information and therefore all follow a threshold strategy.

From now on I assume $\omega < \omega^*$, i.e., the mass of informed agents is not pivotal. Therefore we require

$$q_G(\theta)u(\tilde{r}_2) - u(r_1) > 0 \quad (2.3.6)$$

$$q_B(\theta)u(\tilde{r}_2) - u(r_1) < 0 \quad \text{when} \quad \tilde{n} < \frac{1}{r_1} \quad (2.3.7)$$

Denote $\hat{n} = \frac{R^\eta - r_2}{R^\eta r_1 - r_2}$, which is mass of early withdrawals to make informed agents indifferent between running on the bank and waiting when the future return is expected to be higher than r_2 .

Therefore $n < \hat{n}$ requires

$$\omega > \frac{\lambda + (1 - \lambda) \text{prob}(\theta_i < \theta^*) - \frac{R^\eta - r_2}{R^\eta r_1 - r_2}}{(1 - \lambda) \text{prob}(\theta_i < \theta^*)} \quad (2.3.8)$$

The bank chooses investment and pricing strategies to maximize its expected profit, represented by the following equation:

$$\pi_\eta = \max\{q_\eta(\theta)(1 - nr_1)R^\eta - (1 - n)r_2, 0\} \quad (2.3.9)$$

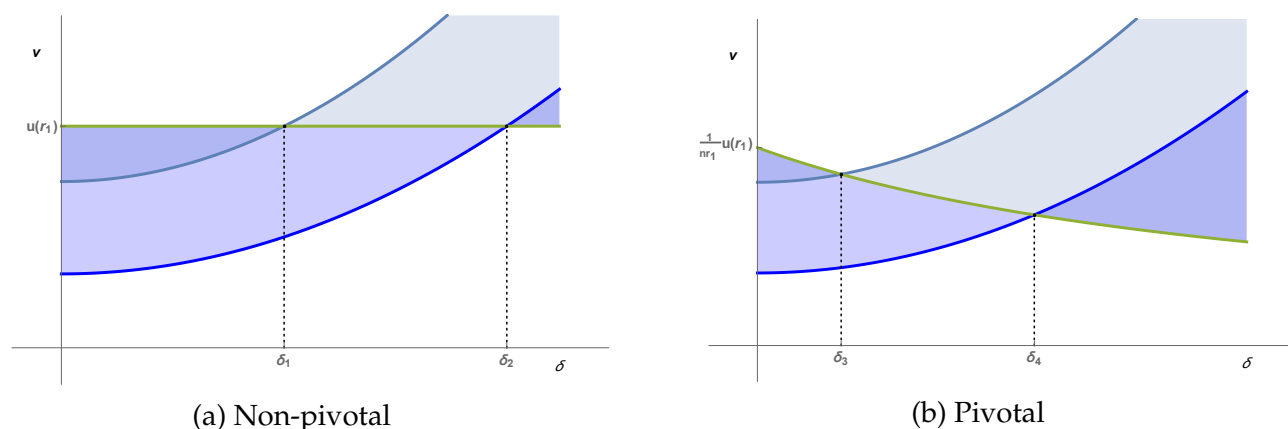
It will be disciplined into choosing good assets if the threat to run strategy of informed agents is credible and important:

$$\pi_G(n = \hat{n}) > \pi_B(n = \hat{n}^l) \quad (2.3.10)$$

2.3.1 Equilibrium Analysis

By using the short-term contract as a monitoring tool, as in [Rey and Stiglitz \(1993\)](#), the bank will find it optimal to hold good assets, even when there is no capital or regulation.

Figure 2.1: Strategies of Agents: Non-pivotal vs. Pivotal Mass of Informed Agents



In period 1, informed depositors learn about the true type of the assets. The market discipline is effective when the bank chooses the good assets for fear of runs of informed

depositors in period 1. An extreme case is when all agents are informed ($\omega = 1$) in period 1. The bank chooses good assets when the discipline is effective. This requires the mass of informed depositors to be large enough. If $n = \lambda + \delta(1 - \lambda) > 1/r_1$, then the run of informed depositors is large enough to induce the bank to liquidate assets.

Moreover, the threat-to-run strategy is credible only if it is incentive-compatible for informed depositors. There is a forbearance region where the informed depositors should run as a discipline tool but is not credible when $\frac{1}{nr_1}u(r_1) < q_B(\theta)u(r_2)$, or $\theta > \hat{\theta}_B$, where $\hat{\theta}_B = q_B^{-1}(\frac{1}{nr_1}u(r_1)/u(r_2))$. When $\theta < \hat{\theta}_B$, the threat-to-run strategy is credible. If the uninformed patient depositors believe that the informed ones follow the discipline strategy and will not run, then they also prefer to wait since $u(r_1) < q_G(\theta)u(r_2)$. Thus the informed depositors will have a stabilizing effect if their deposits is large enough to prevent the bank from liquidation assets.

2.4 Conclusion

Market discipline is not something we should take for granted. This paper studies the liability structure of banks and a novel role of deposit insurance. While the moral hazard problem brought by deposit insurance weakens market discipline, I show that the opposite can be true when the insurance stabilizes uninformed funding and increases the benefits of monitoring through information acquisition. Knowing the bank asset type, informed depositors utilize the demand deposits as a monitoring device and discipline the bank into holding good assets. However, self-fulfilling bank runs initiated by uninformed depositors erodes the future returns, inducing more depositors to forgo information acquisition and act like uninformed depositors. A novel role of deposit insurance emerges from the strategic complementarity between monitoring efforts and stability of uninformed funding. A capped deposit insurance, by stabilizing the retail funding of the bank, restores wholesale depositors' monitoring incentives and benefits market discipline.

Chapter 3

To Float or not to Float? A Model of Money Market Fund Reform

3.1 Introduction

Systemic risks posed by the financial industry are rooted in its role in the creation process of money. Through intermediation between households with savings and firms in need of funding, financial institutions transform relatively long-term investments into demandable debts (bank demand deposits) or equities (money market fund shares), appealing to households who might be hit by a preference shock before the maturity date. It is this function performed by financial intermediaries that helps channel idle funds into productive projects, but at the same time subject them to the vulnerability of liquidity/maturity mismatch.

While demand deposits have been thoroughly explored in classical banking models, money-like claims fabricated by the shadow banking sector are still understudied. This paper approaches the shadow banking money creation process from the perspective of financial contracts and regulation. Traditional banks issue debt contracts embedding a call option, while shadow banks sell equity contracts with a stable redemption price. I show how the equity share contract of money market funds is equivalent with demand deposit contract of commercial banks in providing liquidity achieving the socially optimal allocation.

This paper models money creation in the shadow banking system. Shadow banks (money market funds) issue equity shares to compete with demand deposits offered by traditional banks. With homogeneous households, the two contracts are equivalent in delivering the first-best allocation. Regulatory difference matters when heterogeneous beliefs/asymmetric information are introduced.

The risk-sharing role of demand deposits has been explored in [Diamond and Dybvig \(1983\)](#). The possibility of equity contract to serve the same purpose, nevertheless, receives much less attention in academia. The concept of demand equity and its relationship with demand deposits haven been analyzed in [Jacklin \(1987\)](#), which points out that trading

constraints is essential in making deposit insurance work. Predictions or visions in these papers have become truth, but we lack theoretical framework and tools to anatomize the modern role played by MMMF. The debate over money market fund reform and its reversal calls for a closer examination of the nature of demand equity. [Stiglitz \(2001\)](#), [Brunermeier et al. \(2009\)](#), and [Armour et al. \(2016\)](#) provide insightful and practical thoughts on regulation in the revolving economy.

Several stylized facts about traditional and shadow banking:

1. MMFs emerge as competitors against traditional banks¹. They provide equity share contracts similar to bank debt with on demand redemption, stable NAV and payment functions (defined as the exchange between MMF shares and bank debt in the coexistence equilibrium?)
2. MMF shares are viewed as substitutes for deposits, almost as riskless as insured deposits but with a higher yield. The clientele of MMFs largely overlaps with commercial banks. They suffer similar run concerns which plagued traditional banks before deposit insurance was put in place, and the incentives for MMF shareholders to run is also the same as that for bank depositors: strategic complementarity.
3. The past some thirty years has witnessed how the supply of government securities and a liquid secondary market is developed, and how a full-fledged (and even over developed) financial markets can create private money and reduce financing cost, as well as causing a global financial crisis.
4. It is also during this time that MMF industry experiences considerable growth, especially when money market funds started to provide payment services such as check-writing and wire transfer. It plays an important role in the shadow banking

¹Another extension is when the traditional bankers, instead of being abide by the regulation in the banking industry, move lending activities off their balance sheets through ABCP conduits, which acquire funding from MMFs.

system, with the tipping point coming after Lehman filed bankruptcy in 2008 and the Reserve fund announced breaking the buck.

Besides, regulatory policies matter. While traditional banks are under the heavy regulation in the banking sector, shadow banks are relatively lightly regulated in the securities industry, conforming to information disclosure and investment category restrictions rather than capital requirement and stress-tests. Although not explicitly insured by the Federal Deposit Insurance Corporation (FDIC) like traditional banks, money market funds were fully guaranteed by the government after the Lehman collapse.

Depending on the sophistication of investors, three types of regulation are evaluated and are potentially welfare-improving: (1) capital requirement as discipline tool and buffer ("regulate shadow banks like traditional banks"); (2) information disclosure and floating net asset value to motivate market discipline ("regulate shadow banks like stock market"); and (3) liquidity fees and redemption gates (2016 money market fund reform). Government insurance/guarantee and large-scale asset buying program can also be examined. And empirical implications can be derived and tested using money market fund data.

The model accentuates the importance of FSOC because a statutory power to regulate all financial institutions across industries is vital to identifying shadow banking activities and responding with specific regulatory measures. In addition, the regulation matters not only for financial stability, but also for monetary policies because the migration of activities from traditional to shadow banking system will increase unregulated money creation and weaken the effects of monetary policies.

Given their vital role as a major source of funding in the short-term financing market, the assets side of money market funds has been closely analyzed in various papers, for instance, [Acharya, Schnabl and Suarez \(2013\)](#) on ABCP conduits, [Dang, Wang and Yao \(2014\)](#), [Wang et al. \(2016\)](#), [Allen et al. \(2017\)](#) and [Ruan \(2018\)](#) on China's shadow banking system, and especially, [Chernenko and Sunderam \(2014\)](#), [Schmidt, Timmermann and](#)

Wermers (2016), and Xiao (2018) using MMF data. These analyses take it as given that money market funds are banks without regulation, but still do not investigate why and how they are able to create money-like claims.

Yet there has been little academic research which provides a theoretical framework to analyze the regulatory design on newly-emerged financial sectors. Adrian, Shin et al. (2009) and Gorton and Metrick (2010) propose some principles for regulating the ABCP market and MMFs. These policy recommendations, however, are based on empirical observations and history and lack formal analysis. Gordon and Gandia (2014) examines the run rates of European MMFs during the week when Lehman Brothers filed for bankruptcy. The European MMFs offer a reasonable proxy for the distinction between fixed and floating NAV since they are already issued in two forms, namely “stable NAV” and “accumulating NAV.” The authors find that the stable/accumulating distinction explains none of the cross-sectional variation in the run rate among these funds. Instead, the fund’s portfolio risk and the sponsor capacity to support the fund carry much heavier weights.

The rest of the paper is organized the following way: In Section 3.2, I summarize the institutional background on money market fund (MMF) industry and the 2016 reform. Section 3.3 presents a model with information coordination and derives the equivalence between money market fund shares and demand-deposit contracts. Section 3.4 examines the roles played by the net asset value as a public noisy signal and the equilibrium results under stable and floating NAV regime. Section 3.5 concludes the paper.

3.2 Background on the Money Market Fund Reform

One of the most important lessons learned from the 2007-09 financial crisis is not much different from those learned from other banking crisis and panics: beware of bank runs. And, in the most recent case, beware of shadow bank runs. MMF were a major funding source in the ABCP and repo markets and contribute to the worsening of 2008-09 financial

crisis.

MMF is a type of mutual funds and comprises a percentage in the whole mutual fund industry. The Rule 2a-7 of the Securities and Exchange Commission (SEC) grants critical exemption to money market funds, allowing them to maintain a stable net asset value (NAV) rather than use a floating one as do other mutual funds.

The earliest MMF was founded in the 1972, when Regulation Q governed the traditional commercial banking industry. It provided for common households the access to the treasuries market, which had been restricted to the wealthy due to high investment minimum. Apart from higher yields than deposit interest rates, MMFs also provided check-writing function for shareholders. This payment function is critical for the popularity of MMFs since it made the MMF shares convenient to use and thus a quasi-money ².

In 2008, the Reserve MMF failed to maintain a stable value of \$1 and broke the buck. Panicked shareholders followed by waves of runs on several money market funds. The consequences were so severe that the federal reserve had to intervene and provide full guarantee on the shares.

There have been debates among regulators, scholars and practitioners over how to reform the MMF industry, for instance [Gorton and Metrick \(2010\)](#). The reform bill, passed by the Congress in 2014, has three features: mark-to-market accounting for institutional prime funds ³, redemption gates and liquidity fees. Now MMFs have more tools to cope with runs. The reform dramatically reshaped the MMF industry. The majority of institutional prime funds had converted to treasury funds to avoid the mark-to-market rule before June, 2016, the implementation date.

²Retail money market funds are counted as one of the non-M1 components of M2, along with savings deposits and small time deposits, in the Money Stock and Debt Measures released by the Federal Reserve in the United States. M1 is the more narrowly defined measure which consists of the most liquid forms of money, namely currency and checkable deposits. See <https://www.federalreserve.gov/releases/h6/about.htm> for more details.

³Classification rules of MMFs

3.3 The Baseline Model

3.3.1 The Setup

In this sections, the settings of a model with three periods, $T = 0, 1, 2$, are specified.

There is a continuum of households of unit mass, each endowed with one unit of funds in period 0. The storage technology, which is cost-less and incurs no depreciation, is available to all households.

Following the assumptions made in the seminal [Diamond and Dybvig \(1983\)](#), two types of households are assumed to generate liquidity needs ⁴. A fraction λ of the households are impatient (type- m), and the rest $1 - \lambda$ are patient (type- p). Impatient households requires immediate consumption in period 1, while patient households treat consumption in period 1 and period 2 as perfect substitutes.

Denote c_{i1} and c_{i2} the consumption chosen by type- i household for period 1 and 2, respectively. The preferences are characterized by the following utility function ⁵:

$$U(c_{i1}, c_{i2}) = \begin{cases} U(c_{m1}) & \text{(impatient/type-m)} \\ U(c_{p1} + c_{p2}) & \text{(patient/type-p)} \end{cases}$$

The utility function is assumed to be twice continuously differentiable and $U'(\cdot) > 0$, $U''(\cdot) < 0$. The relative risk aversion is assumed to be constant and larger than one.

Households are identical in period 0, and each have the same probability of being hit by the liquidity shock; that is, revealed to be impatient. In period 1, each household will find out its own type, which will only be revealed to itself (*private information*).

Households will choose the optimal consumption bundle (c_{i1}, c_{i2}) to maximize their own expected utility, $E[U(c_{i1}, c_{i2})]$, given the investment opportunities and technologies

⁴This assumption follows [Diamond and Dybvig \(1983\)](#) and is consistent with the characteristics of depositors and MMF shareholders in comparison to stock market investors.

⁵Assume away the time preference, which will not affect the major conclusions in the model. In appendix I replace the current utility function of patient households with two different forms of $U(c_{p1} + \beta c_{p2})$ or $\beta U(c_{p1} + c_{p2})$, where β is the discount parameter, and the conclusions still go through.

available.

Firms are established by entrepreneurs, who possess skills of doing business but have no endowments⁶. There are numerous entrepreneurs, and each owns a firm with a project in need of one unit of funding.

Projects are long-term and requires two periods to finish. Start with the assumption that all projects are safe and each will yield a gross return Y_s in two periods with certainty. If interrupted in period 1, the fire sale value of the project will be $y_s = (1 - L)Y_s < Y_s$ ⁷, where L measures the extent of liquidation loss.

A benevolent social planner with perfect information on household types and project payoffs will maximize the following social welfare, since there is no aggregate uncertainty.

$$\max_{\{c_{m1}, c_{p1}, c_{p2}\}} \lambda U(c_{m1}) + (1 - \lambda)U(c_{p1} + c_{p2})]$$

subject to the resource constraint

$$(1 - \lambda)c_{p2} = \begin{cases} [1 - \frac{\lambda c_{m1} + (1 - \lambda)c_{p1}}{y_s}]Y_s & \text{if } y_s \geq 1 \\ [1 - \lambda c_{m1} - (1 - \lambda)c_{p1}]Y_s & \text{if } y_s < 1 \end{cases}$$

The optimal consumption bundle $(c_{m1}^*, c_{p1}^*, c_{p2}^*)$ satisfies:

$$U'(c_{m1}^*) = \frac{1}{\max\{y_s, 1\}} Y_s U'(c_{p2}^*)$$

$$c_{p2}^* = \frac{1 - \frac{1}{\max\{y_s, 1\}} \lambda c_{m1}^*}{1 - \lambda} Y_s$$

With the ability to distinguish between patient and impatient households, the social

⁶An equivalent way is to assume insufficient internal funding.

⁷This captures the fact that projects are firm-specific assets; and the price an outsider is willing to pay will be less than the value if the firm could continue the project. It could also be due to lack of bargaining power in a fire sale.

planner will allocate c_{m1}^* to each impatient households in period 1 and c_{p2}^* to each patient households in period 2.

The value of y_s matters for liquidity reserve management. When $y_s \geq 1$, the social planner will invest all funding in productive projects in period 0 and liquidate a fraction of $\lambda c_{m1}^*/y_s$ of the initial investment in period 1. When $y_s < 1$, it will hold a liquidity reserves of λc_{m1}^* from period 0 to 1 and use the reserve to provide consumption for impatient households.

Without loss of generality, assume $y_s = 1$ for simplicity. When $-\frac{cU''(c)}{U'(c)} > 1$, we have $c_{m1}^* > 1$, which is better than the competitive market equilibrium where $c_{m1} = 1$, $c_{m2} = c_{p1} = 0$ and $c_{p2} = Y_s$ ⁸.

Suppose there is a direct financing market where firms can issue equity stock and/or debt contracts to households. In this simple model, the payoffs of projects are riskless and known to all, therefore all firm incomes are pledgeable. Since there is no tax distortion nor bankruptcy cost, the Modigliani-Miller theorem holds, and firms are indifferent between equity and debt financing.

Suppose each household will purchase the stock/bond issued by a firm, which will return Y_s in period 2. Can households potentially do better by trading with one another in a spot market period 1? In period 1, suppose each impatient household will sell its financial claim on the firm at price p . The period 2 return of each claim will be Y_s . By liquidating their own projects, patient households get 1, which is the amount of available funds that can be used to make purchase of claims and obtain a gross return of Y_s/p in period 2. Note that patient households are willing to liquidate own assets and buy claims only if $p < 1$, while impatient households are willing to sell claims rather than to liquidate projects only if $p > 1$. Therefore no trade will happen. The zero trading in the secondary market is because all households will invest in safe projects with no interim payments (no interest nor dividend, since the project needs two periods to generate outputs). Therefore

⁸See [Diamond and Dybvig \(1983\)](#) for detailed analysis and proof.

patient households has no extra funding in hand unless liquidating their own projects.

Will it be profitable to store cash in advance for the opportunity to purchase claims in period 1?

Suppose a fraction λ of a household, instead of investing in safe projects, decides to store the one unit of funding in period 0. In period 1, households each receive a private information regarding their own types. If household who hoards money is impatient, the consumption will be 1. If s/he is patient, s/he can purchase shares from impatient households at a price $p < Y_s$. Note that if $1 < p < Y_s$, the hoarder will be better off by investing in safe projects in period 0. But if $p < 1$, then impatient household will prefer liquidating their assets than trading with the hoarder. In equilibrium, such spot market will not exist, and risk-sharing is not achievable. This is because the financial claim issued by firms does not specify an advance payment in period 1 (i.e., the financial claims are long-term) and is essentially the competitive equilibrium in corporate finance terms.

Now let's revise the direct financing contract a bit so that the firm will distribute a payment (dividend/interest) of d_1 per claim in period 1, and another payment (residual dividend or principal payment plus interest) of $d_2 = (1 - d_1)Y_s$. In this case, the risk-sharing is achieved through trading ex-dividend shares/ex-interest debt contracts in the secondary market as in [Jacklin \(1987\)](#).

In period 1, the supply of ex-dividend equity shares in the secondary market is provided by impatient households at price \mathcal{P} per share. As long as $\mathcal{P} > 0$, impatient households are better off selling the shares. On the demand side, the patient households are willing to pay \mathcal{P} in period 1 to obtain the claim on realization value of projects d_2 in period 2 as long as $\mathcal{P} < d_2$.

The market-clearing condition is $\lambda\mathcal{P} = (1 - \lambda)d_1$, the equilibrium price $\mathcal{P}^* = \frac{1-\lambda}{\lambda}d_1$. Then the total consumption of impatient household in period 1 will be $c_1 = d_1 + \mathcal{P} = \frac{d_1}{\lambda}$, while the consumption of patient household in period 2 is given by $c_2 = \frac{d_2}{1-\lambda} = \frac{1-d_1}{1-\lambda}Y_j$.

By setting $d_1 = \lambda c_1^*$, the direct financing contracts will achieve the socially optimal

allocation. And there are no other equilibria: bank run possibilities are ruled out since only impatient households will sell claims at price \mathcal{P} , which acts as a device to separate different types of agents and reveals information. This is because the equilibrium price of financial claims will fall as a response to excess supply, eliminating strategic complementarity run motive we see in the case of demand deposits.

The obvious advantage of the direct financing market compared to financial intermediaries, which we'll explore in next section, is that firms are not obliged to buying back shares or paying back debts at a fixed price in period 1. The first-best allocation can be delivered through trading among households, and long-term projects will never have to be interrupted.

From now on financial intermediaries will be introduced. Financiers, who are penniless, set up financial intermediaries, which attract savings from households and make loans to entrepreneurs.

Financiers raise money from households either through equity (money market fund shares) or through debt (demand deposits) contracts. Due to the uncertainty of patience types, households favor contracts that provided liquidity in period 1. And only this type of contracts will improve risk-sharing among households compared to the competitive equilibrium.

The money creation process is like this: firms need funding to pay for capital and labor, which are provided by households in exchange for payment. The funding either comes from financial intermediaries, or from financial markets. A firm will choose a certain source of funding which will be accepted by households who might be hit by a preference shock. The means of payment will be accepted in period 0 only if its value in period 1 is also specified in the contracts.

3.3.2 Money Market Funds

Shadow banks also provide funding to firms through loans, but offer equities rather than debts to households⁹. The equity share contract transfers partial ownership of underlying assets (firm loans) to shareholders. The fund pays dividend d_1 per share in period 1 and distribute all residual value $d_2 = (1 - d_1)Y_s$ in period 2 per share.

Trading shadow bank shares in the secondary market is no different from trading firm shares analyzed in the previous section, for the shadow bank is mutually owned by its investors. Given the investors are homogeneous ex ante, the dividend policy will be the same as the one designed by a representative agent.

Here, I'll investigate another version of shadow bank share contracts where the trading in period 1 is conducted directly with the issuer, namely the shadow bank itself.

While the former version allows shareholders to trade ex-dividend equity shares in a secondary market in period 1, the shadow bank in this version is obliged to meet the purchase and redemption demand from investors¹⁰. If the price is purely determined by investors' supply and demand, then the shadow bank merely acts as a central market place, which will not change the results derived in the secondary market trading model. Therefore, the main focus is on the case where the shadow bank offers a fixed price to buy and sell ex-dividend claims, which is exactly how the stable net asset value (NAV) works in the real world.

Suppose the shadow bank offers to sell and buy back ex-dividend shares at a fixed price ϕ in period 1. The immediate result is that there will be zero trading volume in the secondary market, even when it is not forbidden by the shadow bank. This is because

⁹That is, households have (partial) ownership of the shadow bank. This is different from investing in the stock market and acquire (partial) ownership of the firm. Compared with directly participating in the debt market and lending to firms directly, making loans through shadow banks allows for diversification and less monitoring and verification costs. Most importantly, it provides much-needed liquidity through maturity and liquidity transformation.

¹⁰This is different from corporate firms purchasing back shares in the secondary stock market; both prop up the price of securities, but one is not compulsory.

in the secondary market, (1) no buyers will offer a price higher than ϕ since they can do better by purchasing the share directly from the fund; (2) no seller will accept a price lower than ϕ , for selling the share back to the shadow bank is more appealing; (3) given significant searching and transaction costs, sellers and buyers are more willing to directly trade with the shadow bank instead of in the secondary market.

The impatient households will request redemption from the shadow bank and consume $d_1 + \phi$ in period 1, while patient households will each purchase d_1/ϕ shares from the shadow bank ¹¹. The shadow bank can set $\phi = 1$ to replicate the secondary market allocation. The optimal allocation is achieved since all impatient shareholders would like to make redemption with the fund in period 1, while all patient shareholders would like to hold shares and wait until period 2. And when $\phi = 1 = (1 - \lambda)c_1^*$ and $d_1 = \lambda c_1^*$, the demand for redemption $\lambda\phi$ is equal to the demand for new shares $(1 - \lambda)d_1$, the shadow bank does not need any upfront cash or liquidating assets. Essentially, this is equivalent to maintaining a stable NAV fixed at one unit of fiat money per share while distributing $c_1^* - 1 = r_1^*$ dividend shares to early redemption and $c_2^* - 1 = r_2^*$ dividend shares to late redemption.

Again, due to maturity mismatch and liquidity loss, shadow bankers may not be able to maintain the price ϕ . Therefore, the familiar multiple equilibria problem resurfaces: besides the first-best allocation, a bad, self-fulfilling run equilibrium exists. If a shareholder expects all other shareholders to demand early redemption, s/he will also find it optimal to follow the crowd, since the liquidation value of fund assets will not be enough to cover all redemption requests, and the sequential service constraint applies ¹².

Denote γ the number of redemption requests received in period 1. If $\gamma < \frac{1}{d_1 + \phi}$, the status quo can be maintained by liquidating assets to meet the redemption requests. An

¹¹Indeed, the dividends paid by money market funds are usually reinvested in the fund, increasing the number of shares held by the investor.

¹²we can interpret each period represents at least a day, given that money market funds treats intraday redemption requests equally.

investor i will get

$$\tilde{R}(\gamma) = \begin{cases} d_1 + \phi & \text{(early redemption)} \\ \frac{1-\gamma(d_1+\phi)}{1-\gamma} Y_s & \text{(late redemption)} \end{cases}$$

If $\gamma \geq \frac{1}{d_1+\phi}$, the status quo cannot be maintained since the value of redemption requests exceeds the market value of all assets in period 1. The sequential service constraint applies again: all requests received before a certain time period will be redeemed at the stable NAV, while the rest will receive nothing¹³.

The actual payoffs, \tilde{R}_1 and \tilde{R}_2 , respectively, are dependent on γ the number of early redemption in period 1, and γ_i the the number of early redemption requests before investor i . Without loss of generality, assume an investor i will get

$$\tilde{R}(\gamma, \gamma_i) = \begin{cases} d_1 + \phi & \text{(early redemption, } \gamma_i < \frac{1}{d_1+\phi}) \\ 0 & \text{(redemption, } \gamma_i > \frac{1}{d_1+\phi}) \end{cases}$$

This is especially true when the liquidation process of the shadow bank is long and complex, which means that investors have to wait an indefinite number of periods to get their shares redeemed. There will be sufficient incentives for investors to request early redemption¹⁴ Here, another important role played by deposit insurance authority is to provide quick and orderly liquidation.

In the context of MMF shares, the sequential service constraint captures the differ-

¹³It might be an exaggeration to say that late redemption will get nothing back, but it doesn't harm the generality of the model's conclusion. After all, according to documents regarding the Reserve Primary Fund's decision-making the day after the Lehman bankruptcy, the redemption requests the Fund received before 3:00pm on Sep 16, 2008 were still redeemed at the \$1 per share stable NAV, while those requests after 3:00pm were subject to a seven-day redemption suspension permitted by the SEC. Investors lost access to their funds in the account and didn't get the majority of their money back until one month later. The last payment was settled in 2014, six years after the initial suspension. In terms of the loss of liquidity, "receiving nothing" is not quite an overstatement.

¹⁴Similar logics apply to circuit breaker in the stock market, where investors will crowd into selling a stock if they anticipate the trading will soon be suspended, expediting the price fall. This can also be used to analyze the effects of liquidity fees and gates, which may exacerbate the self-fulfilling runs.

ent redemption regimes before and after a critical mass of redemption is demanded. All redemption requests received before a certain time point will be paid the same amount per share. If the total redemption in period 1 is not large enough to destabilize the stable NAV regime, then each investor demanding redemption will receive \$1 dollar per share in period 1, while the rest will receive the pro rata value of realized return of remaining assets in a remote, indefinite period, heavily discounted, which greatly lowers the resale price. The strategic complementarity exists because an investor will be more willing to request redemption if s/he expects more other investors will do the same. The coordination failure leads to a “bad”, self-fulfilling equilibrium, if the interruption of projects is costly enough (low liquidation value in period 1).

3.3.3 Equivalence between MMF Shares and Bank Deposits

3.3.3.1 Demand deposits

The demand deposit is contracted the following way: Each demand deposit contract has a face value of one unit of fiat money. Depositors are free to choose whether to demand early withdrawal in period 1 (early withdrawers), or wait until period 2 (late withdrawers). Bankers promise to pay R_1^* per dollar deposited to early withdrawers, and R_2^* to late withdrawers, where $R_2^* > R_1^* > 1$. Specifically, the payment consists of two parts: interest payment (short-term r_1^* and long-term r_2^*) and principal (par value of the deposit, one unit of fiat money). By definition, $R_1^* = 1 + r_1^*$ and $R_2^* = 1 + r_2^*$. By setting $R_1^* = c_1^*$ and $R_2^* = c_2^*$, the demand deposit delivers the first-best allocation in an incentive-compatible way.

But due to maturity mismatch and liquidation loss, bankers cannot fully honor their promises should all depositors demand early withdrawal. The actual payoffs, \tilde{R}_1 and \tilde{R}_2 , respectively, are dependent on ω the number of early withdrawers in period 1, and ω_i the number of early withdrawers in line before early withdrawer i in period 1.

If $\omega < 1/R_1^*$, the status quo can be maintained by only liquidating some of the bank's

assets to meet the withdrawal need. A depositor i will get

$$\tilde{R}(\omega, \omega_i) = \begin{cases} R_1^* & \text{(early withdrawal)} \\ \frac{(1-\omega R_1^*)Y_s}{1-\omega} & \text{(late withdrawal)} \end{cases}$$

If $\omega \geq 1/R_1^*$, the status quo cannot be maintained since the withdrawal demand exceeds the market value of bank assets in period 1. A depositor i will get

$$\tilde{R}(\omega, \omega_i) = \begin{cases} R_1^* & \text{(early withdrawal, } \omega_i < 1/R_1^*) \\ 0 & \text{(withdrawal, } \omega_i \geq 1/R_1^*) \end{cases}$$

The payoff structure characterizes the sequential service constraint on the financiers' side. The fulfillment of withdrawal requests is on a first-come-first-served basis ¹⁵.

A threshold $\hat{\omega}$ is the number of early withdrawers which will make a patient depositor indifferent between demanding early withdrawal and waiting:

$$\frac{(1 - \hat{\omega} R_1^*) Y_s}{1 - \hat{\omega}} = R_1^*, \quad \hat{\omega} = \frac{Y_s - R_1^*}{(Y_s - 1) R_1^*} < 1$$

As analyzed in [Diamond and Dybvig \(1983\)](#), the good equilibrium can be achieved by setting $R_1^* = c_1^*$ and $R_2^* = \frac{(1-\lambda R_1^*)Y_s}{1-\lambda} = c_2^*$. In the socially optimal equilibrium, only impatient depositors will withdraw money in period 1, and all patient depositors will wait until period 2.

Yet there is a strategic complementarity in demanding early withdrawals. The strategic complementarity leads to multiple equilibria, including a bad equilibrium, namely a self-fulfilling bank run. Expecting other people's early withdrawal will induce a de-

¹⁵A variation incorporates suspension of convertibility and liquidation, where the depositors will line up before the bank opens and the financier will count the number first. It then will decide whether to open business (when $\omega < y_j/r_1^*$) or to declare bankruptcy, liquidate its assets, and allocate the proceeds pro rata to all depositors (when $\omega \geq y_j/r_1^*$). This is more like the case of MMF, where same-day redemption requests are treated equally together after the market closes.

positor to do the same thing. All depositors will withdraw money in period 1 for fear that other people will do the same and nothing would be left in period 2. Even without the inefficiency of excessive risk taking, traditional banks are plagued by runs. Keeping excessive but still fractional reserves to satisfy liquidity needs will not work in times of crisis, and is extremely costly in good times.

In a world with safe projects, only self-fulfilling runs exist. Deposit insurance, which guarantees a payment c_1^* to each early withdrawer, can remedy the case, without having to pay out anything or incurring moral hazard problem. The insured demand deposit contract achieves first-best allocation.

The timeline is the following: In period 0, firms acquire funding from financial intermediaries to pay for capital and labor provided by households, who demand fiat money ("hoarder"), bank deposits ("depositor"), or shadow bank shares ("investor"). Fiat money is used as unit of account of consumption goods.

In period 1, households find out their preference types. Impatient households will withdraw deposits, redeem shares or use fiat money to buy consumption goods right away, while patient ones will decide whether to do the same or wait until next period. The financiers will decide how many projects to liquidate in order to satisfy withdrawal or redemption requests and whether the status quo (full debt repayment or stable NAV) can be maintained. If so, enter into period 2. If not, then the financiers will have to declare bankruptcy/breaking the buck, and liquidate all assets and pay back to the depositors/investors in a pro rata manner.

In period 2, the returns of projects that haven't been liquidated (if any) are realized, and the financiers will pay residual value to depositors/investors who choose to wait until this period.

3.3.3.2 The Equivalence

The capital structure of financial institution doesn't matter in terms of creating liquidity or achieving socially optimal allocation as long as the following conditions are satisfied: (1) Households are identical ex ante, (2) No aggregate uncertainty, (3) The fraction of impatient households are known so optimal dividend policy can be devised, (4) Demand deposits are protected by insurance, (5) Bankruptcy and liquidation are resolved in a costless, timely manner, (6) The secondary market is frictionless. Bank debt contract and fund equity contract are equivalent and trading in the secondary market or directly with the fund also do not matter.

Suppose a financial intermediary issues both demand deposit contracts and equity shares contracts. In period 1, if a household is revealed to be impatient, s/he can either withdraw deposits and receive $R_1^* = c_1^*$, or request redemption/trade the ex-dividend share in the secondary market and get back $d_1 + \phi$ or $d_1 + \sqrt{\quad}$. All are equal to c_1^* . If the household is patient, then s/he will either wait until period 2 and receive $R_2^* = c_2^*$, or use the dividend to purchase more shares directly with the fund or in the secondary market in period 1 and redeem shares in period 2, still getting $(1 + d_1/\phi)d_2 = (1 + d_1/p)d_2 = c_2^*$.

Since the payoff structures are the same, households in period 0 would like to pay identical prices for these contracts. The no-arbitrage condition excludes different pricing of the contracts.

3.4 A Model with Floating Redemption Price

In this section, I'll introduce a risky project and show how heterogeneous beliefs and disagreement on future returns affect the conclusions we have in the previous section.

Suppose the return of the project is risky and its two period return is Y_H with probability $1 - q$ ("good state") and Y_L with probability q ("bad state"). Aggregate uncertainty has been introduced. Suppose $Y_H > Y_s > Y_L$.

Another risk associated with the risky projects lies in its liquidation value in period 1, when the potential outside buyers have figured out the state of the world and the project types but would only like to acquire the superior risky projects at a much deeper discount $(1 - \delta)Y_L$ in the bad state. In the good state, the liquidation value of risky projects will be 1, while that of safe projects is 1 in both states.

Suppose an omnipotent social planner can allocate resources across different states. It aims at smoothing consumption across different types of households and different states of the world.

Then the social planner's problem is

$$\max_{\{c_{m1}^*, c_{p1}^*, c_{p2}^*\}} \lambda U(c_{m1}^*) + (1 - \lambda)U(c_{p1}^* + c_{p2}^*)$$

subject to the resource constraint

$$(1 - \lambda)c_{p2}^* = \begin{cases} [1 - \lambda c_{m1}^* - (1 - \lambda)c_{p1}^*]Y_s & \text{safe projects} \\ [1 - \lambda c_{m1}^* - (1 - \lambda)c_{p1}^*][(1 - q)Y_H + qY_L] & \text{risky projects}^{16} \end{cases}$$

If safe projects are invested, the optimal consumption bundle $(c_{m1}^*, c_{p1}^*, c_{p2}^*)$ satisfies:

$$U'(c_{m1}^*) = Y_s U'(c_{p2}^*)$$

$$c_{p2}^* = \frac{1 - \lambda c_{m1}^*}{1 - \lambda} Y_s$$

If risky projects are invested, the optimal consumption bundle $(c_{m1}^*, c_{p1}^*, c_{p2}^*)$ satisfies:

$$U'(c_{m1}^*) = [(1 - q)Y_H + qY_L]U'(c_{p2}^*)$$

$$c_{p2}^* = \frac{1 - \lambda c_{m1}^*}{1 - \lambda} [(1 - q)Y_H + qY_L]$$

The risky projects are preferred by the omnipotent social planner as long as $(1 - q)Y_H + qY_L > Y_s$. Suppose the inequality above holds. The first-best allocation characterized by (1) investment choice type $j \in \{s, r\}$ and (2) consumption bundle $(c_{m1}^*, c_{p1}^*, c_{p2}^*)$ is $j = r$ and $(c_1^{**}, 0, c_2^{**})$, where

$$U'(c_1^{**}) = [(1 - q)Y_H + qY_L]U'(c_2^{**})$$

$$c_2^{**} = \frac{1 - \lambda c_1^{**}}{1 - \lambda} [(1 - q)Y_H + qY_L]$$

A constrained social planner can only allocate resources intertemporally and across different agents, but not across different states of the economy. The constrained social planner's problem is

$$\max_{\{c_{m1}^G, c_{p1}^G, c_{p2}^G, c_{m1}^B, c_{p1}^B, c_{p2}^B\}} (1 - q)[\lambda U(c_{m1}^G) + (1 - \lambda)U(c_{p1}^G + c_{p2}^G)] + q[\lambda U(c_{m1}^B) + (1 - \lambda)U(c_{p1}^B + c_{p2}^B)]$$

subject to the resource constraint

$$(1 - \lambda)c_{p2}^G = \begin{cases} [1 - \lambda c_{m1}^G - (1 - \lambda)c_{p1}^G]Y_s & \text{safe projects} \\ [1 - \lambda c_{m1}^G - (1 - \lambda)c_{p1}^G]Y_H & \text{risky projects} \end{cases}$$

$$(1 - \lambda)c_{p2}^B = \begin{cases} (1 - \lambda c_{m1}^B + (1 - \lambda)c_{p1}^B)Y_s & \text{safe projects} \\ (1 - \lambda c_{m1}^B + (1 - \lambda)c_{p1}^B)Y_L & \text{risky projects} \end{cases}$$

If safe projects are invested, the optimal consumption bundle $\{c_{m1}^G, c_{p1}^G, c_{p2}^G, c_{m1}^B, c_{p1}^B, c_{p2}^B\}$ satisfies:

$$U'(c_{m1}^G) = Y_s U'(c_{p2}^G), \quad U'(c_{m1}^B) = Y_s U'(c_{p2}^B)$$

$$c_{p2}^G = \frac{1 - \lambda c_{m1}^G}{1 - \lambda} Y_s, \quad c_{p2}^B = \frac{1 - \lambda c_{m1}^B}{1 - \lambda} Y_s$$

Since there is no aggregate uncertainty when investing in safe projects, $c_{m1}^G = c_{m1}^B = c_1^* < c_1^{**}$, and $c_{p2}^G = c_{p2}^B = c_2^* < c_2^{**}$. If risky projects are invested, the optimal consumption bundle $\{c_{m1}^G, c_{p1}^G, c_{p2}^G, c_{m1}^B, c_{p1}^B, c_{p2}^B\}$ satisfies:

$$U'(c_{m1}^G) = Y_H U'(c_{p2}^G), \quad U'(c_{m1}^B) = Y_L U'(c_{p2}^B)$$

$$c_{p2}^G = \frac{1 - \lambda c_{m1}^G}{1 - \lambda} Y_H, \quad c_{p2}^B = \frac{1 - \lambda c_{m1}^B}{1 - \lambda} Y_L$$

Since there are aggregate uncertainties across states, consumption pattern cannot be fully smoothed. If the state of the world is "good" in period 0, the social planner will maintain a liquidity reserve¹⁷ is λc_{m1}^G and invest in risky assets. The optimal consumption bundle is $c_{m1}^G, 0, c_{p2}^G$ which satisfy:

$$U'(c_{m1}^G) = Y_H U'(c_{p2}^G)$$

$$c_{p2}^G = \frac{1 - \lambda c_{m1}^G}{1 - \lambda} Y_H$$

If the state is bad in period 0, then the social planner will invest all funding in safe assets and allocate consumption $c_1^*, 0, c_2^*$. Again, since the social planner can identify the types of households, there will be no runs and no excess liquidation.

Suppose the financial intermediaries cannot write a complete contract conditional on states of the economy. Rather, the financial contracts can only be based on households' action whether the financier has filed for bankruptcy or liquidation protection.

The new timeline is the following: In period 0, nature draws the state of the economy, which is not observable. Financiers specify debt or equity contracts to attract funding from households and choose firm projects to make loans to. In period 1, households' preference types are revealed as private information. Instead of having a consensus on

¹⁷Or by investing in risky assets since their liquidation value is the same as that of safe assets in good state.

the underlying value, a fraction θ of the patient households will investigate the true conditions of the projects and obtain a signal regarding the future payoffs. Start with the assumption that they receive an accurate private signal in period 1 about the true state of the world and corresponding project returns. That is, a fraction $\theta(1 - \lambda)$ households are fully informed about future payoffs. The rest $(1 - \theta)(1 - \lambda)$ of impatient households will remain uninformed because the cost of acquiring information is prohibitively high¹⁸. In period 2, the state of the economy is known to all and the payoffs of the projects are realized.

3.4.1 Insured demand deposits

The new demand deposits is the following: Each demand deposit contract has a face value of one unit of fiat money. Depositors are free to choose whether to demand early withdrawal in period 1 (early withdrawal), or wait until period 2 (late withdrawal). Bankers promise to pay R_1 per dollar deposited to early withdrawal, and the residual value R_2 to late withdrawal, where $R_2 > R_1 > 1$.

The actual payoffs, \tilde{R}_1 and \tilde{R}_2 , respectively, are dependent on ω the number of early withdrawal in period 1, ω_i the the number of early withdrawal in line before early withdrawal i in period 1 and the state of the economy.

When the early withdrawals do not exceed the liquidation value (market price) of bank loans, a depositor i will get

$$\tilde{R}(\omega, \omega_i; \Phi) = \begin{cases} R_1 & \text{(early withdrawal)} \\ \frac{1-\omega R_1}{1-\omega} Y_H & \text{(late withdrawal, good state)} \\ \frac{1-\omega R_1}{1-\omega} Y_L & \text{(late withdrawal, bad state)} \end{cases}$$

¹⁸Many retail depositors/investors do not have the time/expertise/energy/awareness to analyze information, even when the information is disclosed to the public (such as annual reports). They only pay attention to extremely salient features of the financial contracts, such as interest rates (not a wide range of it, but just those advertised by credit card or auto loan companies, such as APY, without knowing the exact meaning and calculation method) and yields, and whether deposits and shares are readily demandable.

When the withdrawal demand exceeds the market value of bank assets in period 1, a depositor i will get

$$\tilde{R}(\omega, \omega_i; \Phi) = \begin{cases} R_1 & \text{(early withdrawal, good state, } \omega_i < 1/R_1) \\ 0 & \text{(late withdrawal, good state } \omega_i \geq 1/R_1) \\ R_1 & \text{(early withdrawal, bad state, } \omega_i < (1-L)Y_L/R_1) \\ 0 & \text{(late withdrawal, bad state } \omega_i \geq (1-L)Y_L/R_1) \end{cases}$$

Without deposit insurance, the strategic complementarity will be more severe if the households find out that the bank invests in risky projects due to the potential liquidation loss.

In a world with insured demand deposits where R_1 is insured for all depositors, only impatient households will demand early withdrawal in the good state, while all patient households will wait until next period since $\frac{1-\lambda R_1}{1-\lambda} Y_H > R_1$.

In the bad state, since $\frac{1-\lambda R_1}{1-\lambda} Y_L < Y_L < 1R_1$, the total amount of depositors demanding early withdrawal will be $\lambda + \theta(1 - \lambda)$ and each will receive R_1 . Instead of the case with sequential service constraint, the deposit insurance guarantees that late-consumers will receive R_1 in the second period, and therefore the rest $(1 - \theta)(1 - \lambda)$ will wait until period 2 expecting a return that can make them at least as good as demanding early withdrawal since they cannot distinguish between different states.

By lending $\theta(1 - \lambda)R_1$ to the bank, the central bank helps the bank to avoid liquidation loss, and replaces the informed households to become entitled to the corresponding payment in period 2. When the state realizes in period 2, the second-period payoff is $R_2^L = \frac{1-\lambda R_1}{1-\lambda} Y_L$. The insurance agency will bear the aggregate risk, paying out the difference $(1 - \lambda)(R_2 - R_2^L)$ so that both the central bank and the uninformed households will get back R_2 per contract. The interest rate earned by the central bank is $\frac{R_2}{R_1} - 1$

By setting $R_1 = c_1^{**}$ and $R_2 = c_2^{**}$, the insured demand deposit can replicate the

omnipotent social planner's allocation.

The only caveat is whether the insurance agency is willing and able to provide the insurance. The total payout of the insurance agency is $v = (1 - \lambda R_1)(1 - q)(Y_H - Y_L)$ in the bad state, and the profit of the bank in the good state is $\pi = (1 - \lambda R_1)Y_H - (1 - \lambda)R_2 = (1 - \lambda R_1)q(Y_H - Y_L)$. Since $(1 - q)\pi = qv$, the insurance is actuarially feasible. By charging a premium π in the good state and paying out v in the bad state, the insurance agency helps deliver the first-best allocation.

One critical implicit assumption is that the insurance agency, public or private, must be able to inject resources into the economy in the bad state. The insurance agency can accumulate liquidity reserve through charging premium in good states, but the reserve may still run short and therefore monetary authority's intervention is needed. Therefore, although it acts as a protection for depositors and banks, the insurance agency itself needs a strong backstop. Only the monetary authority is capable enough to provide liquidity in times of emergency and act as "lender of last resort".

Given the importance of the financial system and the systemic risk presented by financial institutions, the government ultimately will have to step in. This implicit government guarantee or bailout is the root of "too big to fail" headache. To alleviate moral hazard problem posed by systemically important financial institutions, the government must continuously monitor their risk-taking behaviors and performance, as well as reducing the magnitude of the threat they pose to the whole economy. These are exactly why the Financial Stability Oversight Council (FSOC) is created in the wake of the 2007-09 financial crisis.

3.4.2 Secondary market trading

In a world with uninsured equity shares, since the fund cannot anticipate the waves of optimism and pessimism and cannot observe θ in advance, the dividend is contracted the following way: a dividend d_1 per share distributed in period 1, and a residual dividend

d_2 distributed in period 2 with its value depending on the state of the economy.

Denote p the equilibrium price at which impatient households sell their claims in period 1. The informed patient households know the future dividend is $d_2^G = (1 - d_1)Y_H$ in good state and $d_2^B = (1 - d_1)Y_L$ in bad state.

When the true state is good, the informed patient households will act the same way as the uninformed ones, and the market clearing condition is $\lambda p^G = (1 - \lambda)d_1$, $p^G = \frac{1-\lambda}{\lambda}d_1$. The early consumer (impatient households) receive $c_1^G = d_1 + p^G = d_1/\lambda$, while the late consumer (patient households) get $c_2^G = \frac{1}{1-\lambda}d_2^G = \frac{1-d_1}{1-\lambda}Y_H$.

When the true state is bad, the informed households will sell the shares. The market-clearing condition becomes $[\lambda + \theta(1 - \lambda)]p^B = (1 - \theta)(1 - \lambda)d_1$, or $p^B = \frac{(1-\theta)(1-\lambda)}{\lambda+\theta(1-\lambda)}d_1 < \frac{1-\lambda}{\lambda}d_1$. This is because each household owns a share, and if s/he receives a signal pessimistic enough, then there will be more supply than demand, driving down the price.

The consumption of early consumers (impatient and informed patient investors) in period 1 will be $c_1^B = d_1 + p' = \frac{1+\theta(1-\lambda)}{\lambda+\theta(1-\lambda)}d_1 < \frac{1}{\lambda}d_1 = c_1^G$. The pessimism of some patient investors will have a contagious effect through share price, and the pecuniary externality reduces impatient investors' welfare.

In terms of the consumption of late consumers (uninformed patient buyers) in period 2, it is given by $c_2^B = \frac{d_2^B}{(1-\theta)(1-\lambda)} = \frac{1-d_1}{1-\lambda} \frac{Y_L}{1-\theta}$ ¹⁹, which is smaller than the consumption in period 1. The dispersion of beliefs potentially drives down the price of shares exchanged in period 1, enabling uninformed patient investors to purchase more shares, but still consume less than the informed investors.

¹⁹Note that the informed patient households will sell their claims in bad state only if

$$\underbrace{\frac{1-d_1}{1-\lambda}Y_L}_{d_2} < \underbrace{\frac{(1-\theta)(1-\lambda)}{\lambda+\theta(1-\lambda)}d_1}_{p'}$$

$$\underbrace{\frac{1-d_1}{1-\lambda} \frac{Y_L}{1-\theta}}_{c_2} < \underbrace{\frac{1-\lambda}{\lambda+\frac{\theta}{1-\theta}}d_1}_{c_1}$$

The informed investors' driving down the price in bad state has two pecuniary externalities: (1) the impatient investors consume less than when $\theta = 0$; (2) the uninformed patient investors consume more than when $\theta = 0$. These are good spillovers because they improve risk-sharing between preference types under a fixed dividend policy. The interest of informed investors is aligned with the social welfare since they themselves are better off during the process.

The allocation determined in the secondary market involves no run risk and no liquidation loss. But as long as there are sophisticated investors with information about the future state of the economy (i.e., $\theta > 0$), we have $c_1^G > c_1^B$ and $c_2^G > c_2^B, \forall d_1$. In other words, there doesn't exist a dividend policy which can achieve the first-best allocation.

And since investment projects are chosen in period 0, when the state of economy is not observable, the secondary market cannot achieve constrained second-best, either.

Most importantly, the secondary market will disappear if insured bank deposits are introduced. Unsophisticated investors, having anticipated that they will be taken advantage of in the bad state, will choose to deposit their funding in the banks. Sophisticated investors will only participate in the markets if heterogeneous beliefs/disagreement is assumed.

3.4.3 Stable NAV

As a response to the defects of secondary market trading, the shadow bank can intervene and eliminate the pecuniary externalities by participating in the trading.

Suppose the shadow bank offers a fixed price ϕ for equity share purchase and redemption in period 1. Again, the result is that there will be zero trading volume in the secondary market.

When the true state is good, only the impatient investors will redeem their shares, while all patient investors will reinvest the dividend and purchase more. The shadow bank needs to pay out $\lambda\phi$ while receiving $(1 - \lambda)d_1$, which will cancel out with each

other if ϕ is set at $\frac{1-\lambda}{\lambda}d_1$. Each early consumer (impatient investor) will get $c_1^G = d_1 + \phi$ in period 1, while the late consumers will have $c_2^G = \frac{1}{1-\lambda}d_2^G = \frac{1-d_1}{1-\lambda}Y_H$.

When the true state is bad, the impatient investors and informed patient investors will sell their shares back to the fund at ϕ , and uninformed patient investors will use the dividend to buy more shares. The consumption of early consumer (impatient investors and informed patient investors) will be $c_1^B = d_1 + \phi$, while uninformed patient investors will use their dividend to buy $\frac{d_1}{\phi}(1-\theta)(1-\lambda)$ shares and have $(1 + \frac{d_1}{\phi})(1-\theta)(1-\lambda)$ shares in total. Note that the number of newly purchased shares is smaller than the number of redeemed shares $\lambda + \theta(1-\lambda)$ because the market price is fixed by the shadow bank.

The net inflow of shares is

$$\lambda + \theta(1-\lambda) - \frac{d_1}{\phi}(1-\theta)(1-\lambda) = \lambda + [\theta - \frac{d_1}{\phi}(1-\theta)](1-\lambda)$$

The net redemption value is therefore

$$(\lambda + [\theta - \frac{d_1}{\phi}(1-\theta)](1-\lambda))\phi = [\lambda + \theta(1-\lambda)]\phi - (1-\theta)(1-\lambda)d_1$$

Given λ, θ , the shadow bank issuing uninsured equity shares would like to choose d_1 and ϕ to minimize the net redemption value (i.e., net outflow of cash, which requires assets liquidation), which means a high first-period dividend d_1 and a low price ϕ . The lowest feasible value of ϕ is exactly one unit fiat money, since a price higher than one will not minimize the net redemption value, while a price lower than one will make households unwilling to purchase shares in period 0.

When $\phi^* = 1$, the net redemption value will be minimized to zero when

$$d_1^* = \frac{\lambda + \theta(1-\lambda)}{(1-\theta)(1-\lambda)}$$

A feasible d_1 requires $\frac{\lambda+\theta(1-\lambda)}{(1-\theta)(1-\lambda)} < 1$, or $\lambda + \theta(1 - \lambda) < 0.5$. In other words, at least half of the investors should be uninformed and patient to sustain the result. This parameter restriction corresponds to the retail money market fund, where investors are natural persons who seek prudent investments and largely overlap with the clientele of commercial banks.

The ϕ^* and d_1^* achieve market-clearing in the primary market directly operated by the shadow bank. Without relying on liquidity support, the stable NAV helps the shadow bank to avoid liquidating assets, which is vital to surviving adverse economic conditions.

Therefore in the bad state, the early consumers (impatient investors and informed patient investors) will consume $c_1^B = d_1^* + \phi =$, while the late consumers (uninformed patient investors) will consume $c_2^B = \frac{1}{(1-\theta)(1-\lambda)}d_2^B = \frac{1}{(1-\theta)(1-\lambda)}(1 - d_1)Y_L = \frac{1-2\lambda-2\theta(1-\lambda)}{(1-\theta)^2(1-\lambda)^2}Y_L$.

Note that under the optimal dividend policy and stable NAV, the shadow bank needs to pay out λ while receiving $\frac{(1-\theta)\lambda+\theta}{1-\theta} = \lambda + \frac{\theta}{1-\theta} > \lambda$ ²⁰ in the good state. The net value $\frac{\theta}{1-\theta}$ will be stored and redistributed to all remaining shareholders in period 2.

Hence in the good state, the consumption of early consumers (impatient investors) is $c_1^G = d_1^* + \phi^* = \frac{1}{(1-\theta)(1-\lambda)}$ in period 1, while the late consumers (patient investors) will have $c_2^G = \frac{1}{1-\lambda}d_2^G + \frac{\theta}{1-\theta} = \frac{1}{1-\lambda} \frac{1-2\lambda-2\theta(1-\lambda)}{(1-\theta)(1-\lambda)}Y_H + \frac{\theta}{1-\theta}$. The fixed purchase and redemption price avoid liquidation in the bad state, but doesn't help with risk-sharing in the good state.

Let's look back and check whether the uninformed investors will demand early redemption under the fixed NAV through amortized cost method. On book, the total value of the remaining assets in period 2 is changed from $(1 - d_1)Y_H$ to $(1 - d_1)Y_L$, probably through writing down assets, since the state of the economy is know. But should liqui-

²⁰A negative net redemption value, or a positive net new purchase value in period 1, is often not desirable either. This is because the projects have to be invested in period 0 to generate productive income. When more capital is contributed to the open-ended shadow bank, effectively all shares are diluted. This type of dividend policy might not have been approved by the shareholder in period 0, who were identical ex ante. But in this case, the probability of being patient is large enough (over 50%), and it is the patient investors, informed or uninformed, who are the only remaining shareholders of the shadow bank. Therefore the "dilution" is more like a stock split, which will not affect the actual value received.

dation occur, the market value will further drop to $(1 - d_1)(1 - \delta)Y_L$, while δ captures the liquidation loss. With a fixed NAV other than ϕ^* , in order to pay for the net redemption, the shadow bank, which has no liquidity backstop, has to resort to asset liquidation, which will lead to a downward spiral of asset values. Therefore a fixed NAV at $\phi^* = 1$ is the best available tool to reduce net redemption in bad state to zero and avoid asset liquidation.

The amortized-cost net asset value of the shadow bank share in period 1 will be $NAV_{AC} = (1 - d_1)Y_L$, which will be greater than or equal to one if $Y_L \geq \frac{1}{1-d_1}$. The market-to-market net asset value of the shadow bank share after the liquidation in period 1 will be

$$NAV_{MM} = \left[1 - d_1 - \frac{\text{Net redemption value}}{(1 - \delta)Y_L}\right](1 - \delta)Y_L$$

Clearly, in an adverse market environment where the shadow banks desire liquidity the most, $NAV_{AC} > NAV_{MM}$. With the shadow bank itself propping up the price, there will not be information extraction as long as the net asset value is maintained. The uninformed will not panic, and the optimistic will buy the shares at the constant price.

The main difference between traditional banks and MMFs at maintaining a stable price of the contracts lies in the source of liquidity to satisfy the unanticipated early withdrawal/redemption due to mood swings on uncertainties. Traditional banks rely on liquidity backstop provided by deposit insurance and federal reserve at the lender of last resort, which is more efficient and cost-effective in coping with aggregate uncertainties since the central bank controls the money supply. Shadow banks, on the other hand, have to depend on selling assets, often at a discount, to maintain a stable price.

As long as the fraction of patient households receiving negative signals (and therefore requesting redemption) is not very large, the uncertainties regarding future payoff will have no effect on the share price. However, if the fraction is large, more early redemption than λ will either result in interrupting ongoing projects or induce the MMF to

hold more cash reserves. And since the signal is randomly drawn, it is hard to predict the fraction. The only way to make sure is to hold 100% cash reserve, which is the same as the storage and worse than the autarky state (where at least patient households enjoy R units of goods.) Otherwise the large redemption will make it profitable for the rest of the patient household to run as well, making it impossible to sustain the direct redemption and fixed price/NAV regime. The solutions are (1) do not allow direct redemption and let the secondary market work, such as the ETFs; (2) allow direct redemption but with floating NAV to reflect the fair market price; and (3) allow direct redemption and stable NAV only if the liquidation loss is very little.

3.4.4 Amortized-cost vs. Mark-to-market accounting

In this model, since the projects either succeed or fail, the bankruptcy regions for traditional and shadow bankers are the same if they make it to period 2: when the project fails, traditional banker cannot fulfill debt payment, and shadow banker have to liquidate their fund (where nothing remains) An extension allows for more possibilities of investment return/asset price fluctuations, and the shadow banks with no capital buffer will be much more fragile than traditional banks with adequate capital buffer. But if traditional banks are also penniless (infinite leverage), then their fates are the same.. When the project succeeds, traditional bankers will get the residual value, while money market fund managers achieve high return for their investors, keep them in the fund, and earn the management fees ²¹ so that they get residual value rather than a flat-rate fee to remove

²¹Here, the accounting method matters. By using the amortized cost method, the MMFs are able to keep the NAV stable at \$1 per share, while the future expected return from money market debt instruments are linearly distributed as dividends to each share. That is, the MMF pre-allocate the future expected returns, exposing investors who remain in the fund to default or price fluctuation risks. When the underlying investment project is risk-less, then the risk sharing mechanisms of traditional and shadow banks are identical. First-best allocation is achieved in the good equilibrium, and deposit insurance should be put in place to eliminate the bad bank run equilibrium. But when we allow for risky investments and agency problem of bankers, the bankers will take on excessive risk to earn more residual value (in expectation) for traditional bankers and to offer high yields and earn more management fees for shadow bankers. Therefore depositors or investors who choose to wait will be exposed to larger risks, and the risk-sharing feature of these contracts no longer works. Moreover, if households can distinguish between safe and risky projects,

excessive risk-taking incentives if they want to keep stable NAV like bank deposits and provide insurance (when market discipline is costly or there are many unsophisticated investors), or float the NAV to restore market discipline by inducing liquidity-demanding, tail-risk-ignoring (due to inability to distinguish between safe and risky assets) investors to government fund and sophisticated remain in prime fund (better to have a separating equilibrium because in the pooling case, the sophisticated is taking advantage of the naive ones by redeeming shares earlier.)

If the banker chooses to be funded solely by equity, when liquidated, shareholders will be paid first, the same seniority as debt holders when the capital structure is a mixed one. Under the amortized cost accounting method, the net asset value is kept stable at \$1 per share, making the equity share value as constant as debt. The redemption, similar to deposit withdrawals, requires no fees and is subject to the sequential service constraint. Therefore the MMF shares are essentially the same as the demand deposits, although the former is considered as a type of securities and therefore circumvents the capital requirement regulation.

To maintain the contract, MMF managers must ensure a market where the equity shares can be redeemed in period 1, so that impatient households are able to consume goods. The redemption value of each share should not fall below the original purchase price, otherwise households will find the storage more favorable.

Therefore the market in period 1 should satisfy the following conditions:

(1) Equilibrium quantity redeemed is determined by the supply side of the shares, i.e., shareholders requesting redemption;

(2) The demand curve is perfectly elastic, therefore the redemption price is a constant no matter how many shares are redeemed.

or information processing is not costly, market discipline can be restored through the threat of early deposit withdrawals or fund redemption. Otherwise, government intervention is needed. For debt contracts, adding capital buffer (capital requirement) to remove excessive risk-taking incentives and providing deposit insurance to eliminate self-fulfilling runs. For equity contracts, adding another capital buffer with less seniority from the management (similar to hedge funds)

This type of secondary market may often not be readily available. But without the market of redemption, no households would purchase fund shares in period 0. Therefore the fund managers will operate the market and allows redemption so that liquidity is provided.

3.5 Conclusions

This paper models money creation by both traditional and shadow banks and explores optimal regulatory responses. Shadow banks, while performing similar money creation functions as traditional banks, are subject to the regulation on securities (mutual fund) industry rather than the banking industry. Instead of regulated by the central bank and protected by deposit insurance, shadow banks create money through uninsured equity shares and frequently disclose information of their Investment portfolios. A special breed of mutual funds, money market funds are able to do so by investing in liquid money market instruments with highest ratings and thus maintaining a stable net asset value through amortized cost accounting, in contrast to the commonly adopted floating NAV using fair market pricing.

This paper argues that a stable valuation insensitive to market news is critical in producing means of payment/medium of exchange among uninformed investors. Conditions are derived where the information disclosure requirement is equivalent to the classical capital requirement, and where they differ. The regulatory arbitrage weakens the effectiveness of monetary policy, and the private money creation exceeds its socially optimal level. The anticipation of pecuniary externalities due to fire sale and overlapping asset classes through diversification reinforces the run incentives, exacerbating the instability and falling apart of the shadow banking intermediation. Moreover, the accumulation of systemic risk through guarantee and sponsorship links the traditional and shadow banking system, calling for macro-prudential regulation and a central regulator.

Bibliography

Acharya, Viral V, Philipp Schnabl, and Gustavo Suarez. 2013. "Securitization without risk transfer." *Journal of Financial Economics*, 107(3): 515–536.

Adrian, T, H S Shin, et al. 2009. "The shadow banking system: implications for financial regulation." *Financial Stability Review*, , (13): 1–10.

Allen, Franklin, and Douglas Gale. 1994. *Financial innovation and risk sharing*. MIT press.

Allen, Franklin, Yiming Qian, Guoqian Tu, and Frank Yu. 2017. "Entrusted loans: A close look at China's shadow banking system."

Alonso, Irasema. 1996. "On avoiding bank runs." *Journal of Monetary Economics*, 37(1): 73–87.

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

Armour, John, Daniel Awrey, Paul Lyndon Davies, Luca Enriques, Jeffrey N Gordon, Colin P Mayer, and Jennifer Payne. 2016. *Principles of financial regulation*. Oxford University Press.

Begenau, Juliane, and Tim Landvoigt. 2016. "Financial regulation in a quantitative model of the modern banking system."

- Black, Fischer, and Myron Scholes.** 1973. "The pricing of options and corporate liabilities." *Journal of Political Economy*, 637–654.
- Brunnermeier, Markus, Andrew Crockett, Charles AE Goodhart, Avinash Persaud, Hyun Song Shin, et al.** 2009. *The fundamental principles of financial regulation*. Vol. 11, ICMB, Internat. Center for Monetary and Banking Studies.
- Bryant, John.** 1980. "A model of reserves, bank runs, and deposit insurance." *Journal of Banking & Finance*, 4(4): 335–344.
- Chernenko, Sergey, and Adi Sunderam.** 2014. "Frictions in shadow banking: Evidence from the lending behavior of money market mutual funds." *Review of Financial Studies*, 27(6): 1717–1750.
- Dang, Tri Vi, Honglin Wang, and Aidan Yao.** 2014. "Chinese shadow banking: Bank-centric misperceptions."
- Demirgüç-Kunt, Asli, and Edward J Kane.** 2002. "Deposit insurance around the globe: Where does it work?" *Journal of Economic Perspectives*, 175–195.
- Diamond, Douglas W, and Philip H Dybvig.** 1983. "Bank runs, deposit insurance, and liquidity." *Journal of Political Economy*, 91(3): 401–419.
- Donaldson, Jason Roderick, and Giorgia Piacentino.** 2017. "Money runs." *Working paper, Washington University in St. Louis*.
- Dothan, Uri, and Joseph Williams.** 1980. "Banks, bankruptcy, and public regulation." *Journal of Banking & Finance*, 4(1): 65–87.
- Farhi, Emmanuel, and Jean Tirole.** 2017. "Shadow Banking and the Four Pillars of Traditional Financial Intermediation." National Bureau of Economic Research.

- Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny.** 2012. "Neglected risks, financial innovation, and financial fragility." *Journal of Financial Economics*, 104(3): 452–468.
- Gennaioli, Nicola, Andrei Shleifer, and Robert W Vishny.** 2013. "A model of shadow banking." *Journal of Finance*, 68(4): 1331–1363.
- Goldstein, Itay, and Ady Pauzner.** 2005. "Demand–deposit contracts and the probability of bank runs." *Journal of Finance*, 60(3): 1293–1327.
- Gordon, Jeffrey N, and Christopher M Gandia.** 2014. "Money market funds run risk: Will floating net asset value fix the problem." *Colum. Bus. L. Rev.*, 313.
- Gorton, Gary, and Andrew Metrick.** 2010. "Regulating the shadow banking system." *Brookings Papers on Economic Activity*, 2010(2): 261–297.
- Gorton, Gary, and Andrew Metrick.** 2012. "Securitized banking and the run on repo." *Journal of Financial Economics*, 104(3): 425–451.
- Huang, Ji.** 2015. "Banking and Shadow Banking."
- Jacklin, Charles J.** 1987. "Demand deposits, trading restrictions, and risk sharing." *Contractual Arrangements for Intertemporal Trade*, 1: 26–47.
- Jacklin, Charles J, and Sudipto Bhattacharya.** 1988. "Distinguishing panics and information-based bank runs: Welfare and policy implications." *Journal of Political Economy*, 96(3): 568–592.
- Kareken, John H, and Neil Wallace.** 1978. "Deposit insurance and bank regulation: A partial-equilibrium exposition." *Journal of Business*, 413–438.
- Keeley, Michael C.** 1990. "Deposit insurance, risk, and market power in banking." *American Economic Review*, 1183–1200.

- Laeven, Luc, Ross Levine, and Stelios Michalopoulos.** 2015. "Financial innovation and endogenous growth." *Journal of Financial Intermediation*, 24(1): 1–24.
- Merton, Robert C.** 1973a. "An intertemporal capital asset pricing model." *Econometrica*, 867–887.
- Merton, Robert C.** 1973b. "The relationship between put and call option prices: Comment." *Journal of Finance*, 28(1): 183–184.
- Merton, Robert C.** 1977. "An analytic derivation of the cost of deposit insurance and loan guarantees an application of modern option pricing theory." *Journal of Banking & Finance*, 1(1): 3–11.
- Merton, Robert C.** 1995. "Financial innovation and the management and regulation of financial institutions." *Journal of Banking & Finance*, 19(3-4): 461–481.
- Miller, Merton H.** 1986. "Financial innovation: The last twenty years and the next." *Journal of Financial and Quantitative Analysis*, 21(4): 459–471.
- Ordenez, Guillermo.** 2018. "Sustainable Shadow Banking." *American Economic Journal: Macroeconomics*, 2018(10): 33–56.
- Plantin, Guillaume.** 2014. "Shadow banking and bank capital regulation." *Review of Financial Studies*, 28(1): 146–175.
- Pozsar, Zoltan, Tobias Adrian, Adam B Ashcraft, and Haley Boesky.** 2010. "Shadow banking."
- Rey, Patrick, and Joseph E Stiglitz.** 1993. "Short-term contracts as a monitoring device." National Bureau of Economic Research.
- Ronn, Ehud I, and Avinash K Verma.** 1986. "Pricing Risk-Adjusted Deposit Insurance: An Option-Based Model." *Journal of Finance*, 41(4): 871–896.

- Ruan, Tianyue.** 2018. "The Economics of Shadow Banking: Lessons from Surrogate Intermediaries in China."
- Schmidt, Lawrence, Allan Timmermann, and Russ Wermers.** 2016. "Runs on money market mutual funds." *American Economic Review*, 106(9): 2625–57.
- Silber, William L.** 1983. "The process of financial innovation." *American Economic Review*, 73(2): 89–95.
- Stiglitz, Joseph E.** 2001. "Principles of financial regulation: a dynamic portfolio approach." *The World Bank Research Observer*, 16(1): 1–18.
- Stoll, Hans R.** 1969. "The relationship between put and call option prices." *Journal of Finance*, 24(5): 801–824.
- Tufano, Peter.** 1989. "Financial innovation and first-mover advantages." *Journal of Financial Economics*, 25(2): 213–240.
- Wang, Hao, Honglin Wang, Lisheng Wang, and Hao Zhou.** 2016. "Shadow banking: China's dual-track interest rate liberalization."
- Xiao, Kairong.** 2018. "Monetary Transmission through Shadow Banks."

Appendix A

Appendix of Chapter 1

A.1 Interbank Market

In the inter-bank market, there will only be one direction of loans, namely that from patient f -banks to impatient s -banks ¹.

Lemma 10. *There exists an interest rate r_{fs} in the inter-bank market at which f -banks without early redemption requests in period 1 are willing to lend to s -banks with early redemption requests.*

Proof. In period 1, the s -bank facing early redemption (the borrowing bank) will borrow against its asset from the f -bank without early redemption request (the lending bank). The lending bank will sell its safe asset to outside buyers at the price of one, and lend the money to the borrowing bank so as to pay to the impatient household. Thus the s -bank is able to keep the control rights of the assets.

In period 2, the payoff of superior risky asset realizes, and the borrowing bank need to pay back money to the lending bank, which needs to make payments to the f -household. The borrowing bank will receive $(1 - \chi)R_s$ when the project succeeds and $(1 - \chi)R_a$ when the project fails. The lending bank will receive r_{fs} when the project succeeds and R_a when the project fails. The patient household will receive \hat{r}_s^2 when the project succeeds and R_a when the project fails.

The participation constraint of the borrowing bank is

$$(1 - q)[(1 - \chi)R_s - r_{fs}] > 0$$

namely

$$r_{fs} < (1 - \chi)R_s$$

¹ There are four possible inter-bank loan rates, $r_{jj'}$ if a j -bank lends to a j' -bank, where $j, j' \in \{s, f\}$, but only r_{fs} is feasible. Since $r_{sj'} = (1 - \chi_l)R_s$ and $r_{fj'} = R_f$ for $j' \in \{s, f\}$, a patient bank will make zero profits lending to an impatient bank of the same type. And because $r_{sf} = (1 - \chi_l)R_s$ is not affordable to f -banks, impatient safe banks will always choose to liquidate own assets.

The participation constraint of the lending bank is

$$(1 - q)(r_{fs} - \hat{r}_s^2) > R_f - 1$$

namely

$$r_{fs} > \hat{r}_s^2 + \frac{R_f - 1}{1 - q}$$

There exists such r_{fs} if

$$(1 - \chi)R_s > \hat{r}_s^2 + \frac{R_f - 1}{1 - q}$$

$$\chi < 1 - \frac{R_f - q(1 - \chi)R_a}{(1 - q)R_s}$$

which holds for all $\chi < \chi^*$.

Indeed, the inter-bank market works through the r_{fs} .

□

In the inter-bank market, s -banks with early redemption demand will borrow from f -banks with patient households against the superior risky projects in period 1. Essentially, it is as if all impatient households hold safe claims and only safe projects will be liquidated in period 1.

A.2 Enhanced Capital Requirement

Effective shadow bank regulation requires a higher capital requirement than traditional banks. Thus all banks participate in traditional banking (low-cost burdened by high-cost), or create a shadow-shadow banking sector.

If shadow bank regulation is loose than traditional bank regulation, not effective and regulatory arbitrage.

Suppose the regulator expands traditional regulation on safe claims to risky claims.

Since I assume the investment is made after the funding is raised ², the regulator is unable to restrict the investment to safe and superior risky assets as a prerequisite for banks.

Cases are rare when the projects are invested before the funding is raised. And even if the case is true, financial intermediaries may still be able to adjust their portfolios before maturity. So it is reasonable to assume that the regulator can only impose regulation either before the funding is raised or after the projects are invested. This assumption captures the lag of regulation intervention by the regulator in response to financial market activities, even when the regulation rules are specified ex ante.

the possibility of ex ante prohibition in investing in inferior risky projects is excluded because I assume that the government can only acquire information after the projects are invested. That is, it cannot review the information and approve the projects before they are invested. The intervention is afterwards.

Instead, the regulator can base the regulation on the type of financial claims issued.

Assumption 11 (Limited information-processing capability). *When all banks disclose information regarding underlying assets, the regulator can distinguish between safe and risky assets, but can only identify the true types of a fraction η of banks holding risky assets.*

The information set of the regulator is a hybrid version of the banks' and the households', and η measures informativeness of the regulator. When $\eta = 1$, the regulator is informed since it can identify between superior and inferior risky assets, and uninformed when $\eta = 0$.

When the regulator can only identify the types of η banks, no impact on superior risky assets, but banks may have an opportunistic mindset and may try their luck by investing in inferior risky projects.

A capital requirement ensuring "skin in the game" can reduce or even eliminate the excessive risk taking incentives. The banks are able to keep their capital where the state

²The case where the investment is made before the funding is raised is analyzed in Appendix.

is good and financial obligations are met. When the state is bad, the creditors get paid before the shareholders and the capital is seized to compensate for the loss of creditors.

Lemma 12. *Given capital requirement κ and interest rate $r_{\hat{j}}$ ($j = s, f$), banks which issue type- \hat{j} claims will*

1. *prefer type- f over type- i assets if $\kappa > r_{\hat{j}} - \frac{R_f - (1-q)R_i}{q}$;*
2. *prefer type- s over type- f assets if low-cost and prefer type- f over type- s assets if high-cost and $\kappa > r_{\hat{j}} - \frac{R_f - (1-q)(1-\chi)R_s}{q}$;*
3. *prefer type- s assets to type- i if $\chi < \hat{\chi}$ or if $\chi > \hat{\chi}$ and $\kappa > r_{\hat{j}} - (1 - \chi)R_a$*

Proof. (1) Given the interest rate $r_{\hat{j}}$, banks which issue type- \hat{j} claims will invest in type- f rather than type- s if the capital requirement κ satisfies:

$$(1 - q)[(1 - \chi)R_s + \kappa - r_{\hat{j}}] + q\max([(1 - \chi)R_a + \kappa - r_{\hat{j}}], 0) < R_f + \kappa - r_{\hat{j}}$$

when $\kappa < r_{\hat{j}} - (1 - \chi)R_a$, the inequality is equivalent to:

$$\kappa > r_{\hat{j}} - \frac{R_f - (1 - q)(1 - \chi)R_s}{q}$$

which is larger than $r_{\hat{j}} - (1 - \chi)R_a$ for low-cost banks, contradiction.

when $\kappa > r_{\hat{j}} - (1 - \chi)R_a$, the inequality is equivalent to:

$$(1 - \chi)[(1 - q)R_s + qR_a] < R_f$$

which does not hold for low-cost banks, contradiction.

Thus the condition will never hold for low-cost banks, but for high-cost banks, the condition is equivalent to $\kappa > r_{\hat{j}} - \frac{R_f - (1-q)(1-\chi)R_s}{q}$.

(2) Given the type- \hat{j} ($j = s, f$) claims, banks which issue type- \hat{j} claims will invest in type- f rather than type- i if the capital κ satisfies:

$$R_f + \kappa - r_{\hat{j}} > (1 - q)[R_i + \kappa - r_{\hat{j}}] + q \max[(\kappa - r_{\hat{j}}), 0]$$

when $\kappa < r_{\hat{j}}$,

$$\kappa > r_{\hat{j}} - \underbrace{\frac{R_f - (1 - q)R_i}{q}}_{>0}$$

when $\kappa > r_{\hat{j}}$, the inequality is equivalent to

$$(1 - q)R_i < R_f$$

which holds all the time.

(3) Given the interest rate $r_{\hat{j}}$, banks which issue type- \hat{j} claims will invest in type- s rather than type- i if the capital requirement κ satisfies:

$$(1 - q)[(1 - \chi)R_s + \kappa - r_{\hat{j}}] + q \max([(1 - \chi)R_a + \kappa - r_{\hat{j}}], 0) > (1 - q)[R_i + \kappa - r_{\hat{j}}] + q \max([\kappa - r_{\hat{j}}], 0)$$

if $\kappa < r_{\hat{j}} - (1 - \chi)R_a$, the inequality is equivalent to:

$$(1 - \chi)R_s > R_i$$

which holds for $\chi < \hat{\chi}$.

if $r_{\hat{j}} - (1 - \chi)R_a < \kappa < r_{\hat{j}}$, the inequality is equivalent to:

$$\kappa > r_{\hat{j}} - \frac{(1 - \chi)[(1 - q)R_s + qR_a] - (1 - q)R_i}{q}$$

which is smaller than $r_{\hat{j}} - (1 - \chi)R_a$ and is not binding.

if $\kappa > r_{\hat{j}}$, the inequality is equivalent to:

$$(1 - \chi)[(1 - q)R_s + qR_a] > (1 - q)R_i$$

which holds all the time for all banks if assume $(1 - \chi)[(1 - q)R_s + qR_a] > 1$.

when $\kappa > r_{\hat{j}} - (1 - \chi)R_a$, the inequality is equivalent to:

$$(1 - \chi)[(1 - q)R_s + qR_a] < R_f$$

which does not hold for low-cost banks, contradiction.

Thus the condition will never hold for low-cost banks, but for high-cost banks, the condition is equivalent to $\kappa > r_{\hat{j}} - \frac{R_f - (1 - q)(1 - \chi)R_s}{q}$. \square

Direct revelation mechanism can be achieved by setting capital requirements $\kappa \geq \kappa_{TB}$ on traditional banks and $\kappa \geq \kappa_{SB}$ and shadow banks, where

$$\kappa_{TB} = r_{TB} - \frac{R_f - (1 - q)R_i}{q}, \quad \kappa_{SB} = r_{SB} - \frac{R_f - (1 - q)R_i}{q}$$

Note that low-cost banks with $\chi < \hat{\chi}$ will always invest in superior risky assets. That is, the capital requirement cannot prevent banks from financing superior risky assets through safe claims.

In order for the deposit insurance or implicit guarantee not being abused, need to require $\kappa_{TB} > r_{TB} - q(1 - \chi_l)R_a$. To ensure the financial claims are indeed safe from the perspectives of households, the capital requirement should be able to deliver full payment even when the underlying projects fail, namely $\kappa > r_{TB} - (1 - \chi_l)R_a$, which is satisfied by κ_{TB} .

The capital requirement in shadow banking sector plays a dual role: it prevents excessive risk-taking of low-cost banks when $\chi_l > \hat{\chi}$, as well as preventing regulatory arbitrage by inferior risky assets. But when $\chi_l > \hat{\chi}$, low-cost banks always prefer the superior

risky projects, and the high capital requirement in shadow banking only serves to prevent regulatory arbitrage, which creates a barrier to entry for innovative low-cost banks when banks' capital is scarce.

When $\kappa > \kappa_{SB}$, high-cost banks prefer safe assets over risky ones and will operate in traditional banking. Low-cost banks, however, will also operate in traditional banking for the lower interest rate.

In equilibrium, all bank participate in traditional banking. The interest rate is not changed and the first-best is achieved under the capital requirement, which prevents high-cost banks from excessive risk-taking while helping stabilize the interest rate when low-cost banks invest in superior risky assets.

When $\kappa < \kappa_{SB}$, banks will not be able to participate in shadow banking, which is not a binding constraint since no banks will raise funding in the regulated shadow banking. In the traditional banking sector, high-cost banks will invest in safe assets and low-cost banks will still innovate and invest in superior risky assets.

When $\kappa < \kappa_{TB}$, banks will not be able to participate in traditional banking. Beneficial innovation is sacrificed for preventing excessive risk taking, and all wealth are stored.

Condition 13 (No-regulatory-arbitrage condition). *To prevent regulatory arbitrage, the capital requirement in the shadow banking sector should not be less than that in the traditional banking sector:*

$$\kappa_{SB} \geq \kappa_{TB}$$

Otherwise, banks will choose to issue risky claims and we're back to laissez-faire banking.

By requiring a risk-based capital requirement $\kappa = r_{\hat{j}} - \frac{R_f - (1-q)R_i}{q}$, banks issuing \hat{j} -claims will only invest in safe assets.

Appendix B

Appendix of Chapter 2

B.1 Model with a Wealthy Depositor

Suppose there is a wealthy depositor endowed with $W > 1$ units of wealth in period 0.

Particularly, I assume the wealthy depositor will receive an income of I in period 1. This allows for additional liquidity which can be used to support banks facing runs of common depositors. In this sense, the wealthy depositor plays the role similar to a central bank as "lender of last resort". The benchmark case where the wealthy depositor has unconstrained income gives us the usual result as in the social planners case. However, the focus here is on whether the market itself could survive panics without relying on government backstop. This is emphasized later when I add constraints on liquidity/capital held by the wealthy depositor in period 1 to accentuate the difference. Normally, the state is assumed to have exuberant power in terms of providing liquidity through various monetary and fiscal policies, but not the private market.

The wealthy depositors, however, are always patient and risk-neutral (or use convex utility function), indifferent between consumption in period 1 and in period 2. Rather than risk-sharing, the wealthy depositor uses demand-deposit contract as monitoring device.

Wealthy depositors have incentives to deviate since $\alpha w_h > c$. If the project is bad, running on the bank is the weakly dominant strategy. If the funding from the wealthy depositor is large enough, then this will exert monitoring effort on the bank and lead to effective market discipline. The uninformed depositors are simply free-riders. If they can observe the wealthy depositor's withdrawal behavior, then it is equivalent to the case where the private signal is made public. The signal becomes noisy if the uninformed depositors cannot tell whether the wealthy depositor withdraws due to bad signal or its own liquidity preference.

If the project is good, then the strategy depends on whether the wealth is constrained: if other uninformed depositors choose to run, the wealthy depositor will stay in the bank

as long as the run does not threaten the solvency of the bank. Moreover, it will lend more to the bank to satisfy the liquidity needs. Alternatively, it will buy the deposit contracts from panicked uninformed depositors. In this sense, the wealthy depositor has an anchoring effect in the bank's funding structure.

However, if the funding from the wealthy depositor is relatively small, then there is a strategic complementarity among informed and uninformed depositors: the wealthy depositors may run even when the private signal is good, since the potential run from uninformed depositors will threaten the solvency of the bank due to fire sale discounts. In this case, the wealthy depositors will not acquire private information to begin with. All depositors are uninformed; no market discipline, and the banking system is not viable if there is no regulation such as capital requirement to curb the excessive risk-taking incentives.

B.2 Model with Separated Contracts

In this section I examine the possibility of the bank offering two separate contracts to common and wealthy depositors, namely different rates for retail and wholesale funding.

In the model, there are two types of creditors: retail depositors and wholesale lenders. Banks can choose from two investment portfolios different in risk levels, and there is a fixed cost to monitor banks.

The wholesale lender has more money to lend and is more financially sophisticated, which brings in two effects, depending on the overall economic situation. In good times, the wholesale lender is not liquidity-constrained and can meet the borrowing demand from the bank, obtaining the portfolio as collateral. Even if when all retail depositors run on the bank, the bank can always fill in the liquidity gap by borrowing from the wholesale lender, which disincentivizes the retail depositors from bank runs to begin with. This is the **anchor effect**, which increases the stability of the bank.

In bad times, however, the bank's portfolio worth less due to price fall of assets, and

the wholesale lender may not be willing to roll over the debt (or “flight to quality”). In this case, the bank has to liquidate some assets in the portfolio in order to pay back the debts, which may exacerbate the fire sale problem. Retail depositors will also join the force and demand early withdrawal. In the end, the whole portfolio is liquidated, the bank goes bankrupt, and the fire sale problem is even more severe in the market, which would result in more bank failures.

When deposit insurance is put in place and all retail depositors are fully insured, market discipline incentive will not be reduced because these depositors wouldn’t exert efforts to monitor the bank even when not protected by the insurance. The major benefit of deposit insurance is less assets liquidation needed in bad times, therefore alleviating the fire sale problem. The cost, however, is the loss that will be incurred to the insurance fund if the bank still fails.

When the deposit insurance limit is increased to cover all wholesale debts, then market discipline incentive will be eliminated, and the insurance fund faces an even larger payout when the bank fails. The benefit, nevertheless, is that the bank’s ability to roll over its debt no longer depends on the marked-to-market value of the collateral it can provide. In other words, the bank remains immune to the financial contagion through fire sale of assets by other institutions.

If they choose to wait until the last period when the assets mature, namely, they agree to rollover the one-period debt contract with the bank, then they will get $r_D^2 D$ and $r_U^2 U$, respectively, if the project succeeds, and get 0 otherwise. If the households choose not to rollover the debt, then the bank can avoid liquidating assets by borrowing more from the wholesale lender, if the wholesale borrowing rate is low enough.

The timeline is the following:

1. In period 0, retail depositors enter into the demand deposit contract and the wholesale lender negotiates a one-period debt contract with the bank. The banker will choose one portfolio and invest $A = D + U + E - R$ while holding R amount of

reserves. The choice of portfolio will not be revealed to both types of creditors, but they can choose whether to exert efforts to discover the truth, which takes one period to work.

2. In period 1, wholesale lender will decide whether to rollover the debt and retail depositors will decide whether to demand early withdrawal. If they have exerted monitoring efforts, then they will know which portfolio the bank has chosen and invested in. The bank will liquidate some or all of its assets in the portfolio at the market price to meet the liquidity demand, if any.
3. In period 2, the remaining assets (if any) in the portfolio mature. If the underlying projects are successful, then the debt contracts are fulfilled and the banker gets the residual value as profits. If not, then the banker gets nothing, and the reserve will be split among creditors according to some predetermined rules (e.g., seniority).

Run region. Given the bank's choice of portfolio,

(1) if the wholesale lender agrees to roll over the debt and retail depositors do not demand early withdrawal, the expected utility of the wholesale lender is $p_H u(r_{U2}U) + (1 - p_H)u(\lambda_U R)$, and the expected utility of the retail depositors is $p_H u(r_{D2}D) + (1 - p_H)u(\lambda_D R)$.

(2) if the retail depositors all demand early withdrawal and the wholesale lender agrees to lend more to meet the liquidity needs, then each depositor gets $r_{D1}D$ while the expected utility of the wholesale lender is $p_H u(r_{U2}U + r_{U1}r_{D1}D) + (1 - p_H)u(R)$.

(3) if the wholesale lender declines to rollover the debt while the retail depositors do not demand early withdrawal, then the bank has to liquidate some assets in order to meet the liquidity assets, $(1 - d)L = r_{U1}U$, and has $A - L$ remaining assets. The expected utility of depositors is $p_H u(r_{D2}D) + (1 - p_H)u(R)$ if the remaining assets satisfies $r_H(A - L) + R > r_{D2}D$. If $r_H(A - L) + R < r_{D2}D$, then the expected utility of depositors is $p_H u(r_H(A - L) + R) + (1 - p_H)u(R)$.

(4) if both types of creditors decide to quit in period 1, then all assets are liquidated. The total value of the bank is $(1 - d)A + R$, which is allocated between creditors and the banker. The banker's equity can be wiped out if d is large enough. Suppose $(1 - d)A + R > r_{U1}U + r_{D1}D$.

Given the wholesale lender agrees to roll over the debt, the retail depositors will run if $u(r_{D1}D) > p_H u(r_{D2}D) + (1 - p_H)u(\lambda_D R)$.

Given the wholesale lender declines to roll over the debt and $r_H(A - L) + R > r_{D2}D$, the retail depositors will run if $u(r_{D1}D) > p_H u(r_{D2}D) + (1 - p_H)u(R)$. That is, if U is small enough, the run region of retail depositors is larger if the wholesale lender doesn't run. This is because wholesale funding and retail funding are competing for the reserves when the investment portfolio fails if both decide to wait until the last period.

Given the wholesale lender declines to roll over the debt and $r_H(A - L) + R < r_{D2}D$, the retail depositors will run if $u(r_{D1}D) > p_H u(r_H(A - L) + R) + (1 - p_H)u(R)$. That is, if U is large enough, the run region of retail depositors is smaller if the wholesale lender doesn't run. This is because the liquidity demanded by wholesale funding in period 1 is so large that long-term investment will be deteriorated to a large extent. Therefore the wholesale funding can stabilize retail funding when it doesn't run, and will destabilize retail funding when it runs.

Given the retail depositors do not demand early withdrawal, the wholesale lender will decline to roll over the debt if $u(r_{U1}U) > p_H u(r_{U2}U) + (1 - p_H)u(\lambda_U R)$.

Given the retail depositors demand early withdrawal, the wholesale lender will decline to roll over the debt if $u(r_{U1}U) > p_H u(r_{U2}U + r_{U1}r_{D1}D) + (1 - p_H)u(R)$ and $(1 - d)A + R > r_{U1}U + r_{D1}D$.

The run region of wholesale lender will be larger if the retail depositors don't run. This is because when the retail depositors don't run, they will be competing for the reserve, and the wholesale lender will lose the opportunity to lend more to the bank and potentially get higher expected utility.

Bank runs. In addition, suppose there is a systemic risk with probability s at time 1. Then the bank has to liquidate some of its assets in order to pay the creditors¹. Put differently, the bank is no longer able to rollover its short-term debt.

In normal times, the wholesale lender is not liquidity constrained, therefore will agree to extend the lending if the less risky portfolio is chosen. In times of crisis, however, liquidity in the financial market evaporates, and the wholesale lender will also decide not to rollover the debt.

Therefore in normal times, the wholesale lender acts like an anchor lender which boosts small creditors' confidence in the bank and enhances stability. However, when there is a liquidity shock, the wholesale lender will stop lending and become a financial destabilizer, since it will induce small depositors to stop rolling over the debt.

In normal times (with probability of $1 - s$), other lenders' decision depends on wholesale lender's decision. The wholesale lender will roll over the debt only if it finds out the bank has chosen the less risky portfolio. Therefore the bank will choose the less risky portfolio. In times of crisis (with probability of s), the wholesale lender will not roll over the debt and require the bank to fulfill the payment of $r_{U1}U$, and retail depositors will also demand early withdrawal of $r_{D1}D$. All assets have to be liquidated at a low fire sale price (or high discount rate when used as collateral to get new funding), and since $[1 - \delta(A)]A < r_{D1}D + r_{U1}U$, the banker's equity is wiped out and losses are incurred to creditors.

The banker's expected utility is

$$(1 - s)p_L u(r_L A - r_{U2}U - r_{D2}D) \tag{B.2.1}$$

¹It could be the case where the bank's balance sheet shrinks by the factor δ , either through loss from transactions with the failed systemic institution, or through falling asset price resulting from a portfolio similar to that held by the failed systemic institution and fire sale price. It could also result from the investor/creditor side due to evaporated liquidity and flight to quality: the repo or ABCP market freezes and creditors are no longer willing to rollover their debt. In this paper, the behavior is modeled as creditors "run" on the bank, and the mechanism is basically the same as that in a model of debt rollover failure.

$$(1 - s)p_L u[r_L E + (r_L - r_{U2})U + (r_L - r_{D2})D] \quad (\text{B.2.2})$$

which is assumed to be greater than $u(E)$; i.e., the participation constraint is satisfied.

If the bank chooses the more risky portfolio, then even in normal times, debt will not be rolled over and bank profit is zero.

Banks are good at managing risky financial assets, and there will be social welfare loss when these assets are liquidated before maturity (financial disintermediation). In other words, investment technology is available only to bankers, who are financially more sophisticated. Depositors/creditors only have access to risk-free assets (or storage technology), of which the return is certain and given.

Full insurance. Now we assume that both types of debts are fully insured. Then there will be no bank runs and no market discipline. Rather, a higher insurance premium is charged.

If the banker chooses the less risky portfolio, her expected utility is

$$p_L u[r_L A - r_{U2}U - r_{D2}D] \quad (\text{B.2.3})$$

If the banker chooses the more risky portfolio, her expected utility is

$$p_H u[r_H(A - a) - r_{U2}U - r_{D2}D] \quad (\text{B.2.4})$$

where a punishing insurance premium a is charged on the bank.

The banker will prefer the less risky portfolio if

$$p_L u[r_L A - r_{U2}U - r_{D2}D] > p_H u[r_H(A - a) - r_{U2}U - r_{D2}D] \quad (\text{B.2.5})$$

Compared with the case where only deposits are fully insured:

$$p_L u(r_L A - r_{U2} U - r_{D2} D) > \frac{p_H}{1-s} u[r_H(A-L) - r_{D2} D] - \frac{s}{1-s} p_L u[r_L(A-L) - r_{D2} D] \quad (\text{B.2.6})$$

The optimum risk premium is pinned down by solving

$$p_H u[r_H(A-a) - r_{U2} U - r_{D2} D] = \frac{p_H}{1-s} u[r_H(A-L) - r_{D2} D] - \frac{s}{1-s} p_L u[r_L(A-L) - r_{D2} D] \quad (\text{B.2.7})$$

In this case, market discipline is preserved by charging a punishable insurance premium on the bank. The social welfare is higher because the investment is not interrupted.

The question is whether the centralized monitoring can work better than decentralized market discipline. It could be yes, since there are high-cost and free-rider problems with market discipline and even large creditors can shift the risk by purchasing CDS to avoid costly monitoring. It could be no, since it's never easy to find the right level of risk premium, and large centralized agency is always associated with the problem of inefficiency.

The debate, ultimately, is about the advantages and disadvantages of market and government, the invisible hand and the visible hand.

It seems that when there is financial friction, prominently in the systemic crisis (fire sale, pecuniary externality, financial contagion), government intervention and centralized planning is more relevant and efficient.

B.3 A Deposit Insurance Fund facing Tradeoffs

This model goes back to the setting where there is no excessive risk-taking activities and no information acquisition regarding the bank assets. The payoff is solely determined by the state of the nature, and self-fulfilling bank runs coexist with fundamental-driven bankruptcy.

Here, the capped deposit insurance is rationalized by a tradeoff between preventing bank runs and reducing payout costs.

B.3.1 The Environment

We follow the basic settings in the seminal Diamond and Dybvig (1983). In this three-period model ($T = 0, 1, 2$), there is a single consumption good, and a continuum of *ex ante* homogeneous agents of measure one in this model.

There are two *ex post* types of agents: impatient and patient. A fraction $\lambda \in (0, 1)$ ² of the agents are impatient, which is common knowledge. The probability of being impatient, conditional on λ , is equal and independent for each agent.

Impatient agents only care about their consumption in $T = 1$ (denoted as c_m) and their utility function is $u(c_m)$, while patient agents can consume in both periods, c_{p1} and c_{p2} , and only care about the total amount of consumption with a utility function of $u(c_{p1} + c_{p2})$ ³, where u is twice continuously differentiable, increasing and strictly concave.

The agents have access to an investment technology, which returns R ⁴ units of output in $T = 2$ for each unit of input invested in $T = 0$ ⁵. The technology also allows the consumers to get their initial investment when the project is interrupted in $T = 1$ ⁶.

Now let's discuss the case where fundamental bank runs are introduced. There is a positive probability, q , with which the economy will transit from the solvent state to the insolvent state. And we also need to check that banks are viable, *ex ante*.

²Here, we assume λ is fixed. Diamond and Dybvig (1983) also discussed the case allowing λ to be random (aggregate uncertainty). Later in the extensions part, we will relax this assumption and assume a random proportion.

³For impatient agents, the time-preference parameter, ρ , is simply zero. The patient agents have a positive ρ , but the magnitude is irrelevant for our current research. Without loss of generality, we assume $\rho = 1$ and focus the research on R .

⁴In Diamond-Dybvig model, they assume the long-run return is risk-free, while in Goldstein-Pauzner model (2005), the investment project returns R with probability $p \in (0, 1)$ and returns 0 otherwise.

⁵We will relax the condition and discuss the case where liquidation technology returns smaller than one, and the case where the fundamentals are so good that the short-run return also improves.

⁶Therefore, compared to the storage technology, the investment technology is as good if the agent is impatient, and is strictly better if the agent is patient.

Now let's first consider a simple model where there are only two possible states of the world: H and L , corresponding to R_H and R_L , respectively.

In state H , the investment project returns R_H , while in state L , the investment project returns R_L , and $R_H > 1 > R_L$ ⁷. Therefore when $R = R_L$, it is efficient to liquidate the assets, and both types of agents will run on the banks at $t = 1$.

The timing of events is the following:

- In period 0, nature draws each agent's type from independent and identical distributions. This information, however, is not disclosed to each agent until next period. Agents each will receive one unit of endowments and make investments. Note that with the costless liquidation technology and a risk-free return larger than one, agents will invest all of their endowments in projects in this period.
- In period 1, the type of each agent is revealed as a private information. Impatient agents will interrupt the investment and make consumption. Patient agents will choose whether to consume now, or wait until next period.
- In period 2, the return of remaining investment, if any, is realized and patient agents will consume all goods available, if they choose to wait in the previous period.

The sequence of actions is the following: nature moves first and draws the state, H with probability $1 - q$ and L with probability q . For each realized state, there are three periods, 0, 1 and 2, the same as the setting in Diamond-Dybvig model. *Ex ante* homogeneous agents make investment decisions in period 0. In period 1, the true state of the world is revealed, as well as each agent's type. The former is a common knowledge, while

⁷We can also introduce risks into the return. One way is to assume that the investment returns R_H with a positive probability of success in H state, and returns R_L with the same probability of success in L state. The probability of success among different states is the same, and can either be a constant or a function of bank's monitoring effort. Another way to capture differences between these two states is to assume that given a certain level of return, the probability of success is higher in H state, which will not make fundamental difference.

the latter is a private information. The impatient agents will demand early withdrawal, while the patient agent chooses whether to wait or run.

In an economy without banks offering demand deposit contract, agents first invest their endowments in projects at period 0. In period 1, the state of the world is revealed. If $R = R_H$, the state is solvent, then we're in the same autarky case as in the Diamond-Dybvig model. If the return is revealed to be $R_L < 1$, however, then both types of agents will liquidate their assets.

The social welfare is

$$\begin{aligned} W'_{autarky} &= (1 - q)[\lambda u(1) + (1 - \lambda)u(R_H)] + qu(1) \\ &= \lambda u(1) + (1 - \lambda) \underbrace{[(1 - q)u(R_H) + qu(1)]}_{E u(R)} \end{aligned} \quad (\text{B.3.1})$$

Here, our focus is on the case when $0 < q < 1$ ⁸

The optimal risk sharing problem is

$$\max (1 - q)[\lambda u(c_m) + (1 - \lambda)u(c_{p2})] + qu(1) \quad (\text{B.3.2})$$

subject to the resource constraint

$$\lambda c_m + \frac{(1 - \lambda)c_{p2}}{R_H} = 1 \quad (\text{B.3.3})$$

This is equivalent to maximizing $\lambda u(c_m) + (1 - \lambda)u(c_{p2})$. Therefore we have

$$u'(c_m^*) = R_H u'(c_{p2}^*) \quad (\text{B.3.4})$$

⁸When $q = 1$, the state of the world must be L . Since R_L is so low, early liquidation becomes desirable. When $q = 0$, we go back to the one-state case in the previous section except for the fact that the return to investment is R_H rather than R_{DD} . *Ex ante*, the bank can set $r_1 = c_m^*$ where $u'(c_m^*) = R u'(c_{p2}^*)$ and $c_{p2}^* = (1 - \lambda c_1^*)R / (1 - \lambda)$ to reach optimal risk sharing.

where

$$c_{p2}^* = \frac{(1 - \lambda c_m) R_H}{1 - \lambda} \quad (\text{B.3.5})$$

B.3.2 Equilibrium with Banks

Denote f_i the number of depositors standing in front of depositor i in the line and V the total number of depositors in the line demanding early withdrawal.

Note that bank run is defined as the case where patient depositors also demand early withdrawal, i.e., the "bad" equilibrium described in the Diamond-Dybvig model. Note that the minimum value of V is λ , the number of impatient agents. Therefore an implied constraint on the deposit insurance premium is $\frac{1-a}{r_1} \geq \lambda$, or $a \leq 1 - \lambda r_1$, since the λ impatient agents will always withdraw money in period $T = 1$ (otherwise V will always be in the second case and the insurance fund always needs to pay out money, making banks not desirable in the first place). This poses an upper bound to the insurance premium.

Note that since impatient agents will always withdraw in $T = 1$, whether there will be bank runs depends on whether the patient agents will demand early withdrawal and if so, how many.

When there are V depositors lining up in front of the bank counter and demanding early withdrawal, the rest of patient agents will compare the return of waiting until period $T = 2$ and that of demanding early withdrawal in period $T = 1$, namely r_2 and r_1 . Note that r_1 is a fixed amount promised in the demand deposit contract, while r_2 is the residual value of banks divided by the number of depositors waiting until the last period.

When there is no deposit insurance, the payoff in period $T = 2$ is

$$r_2 = \begin{cases} \frac{(1-r_1V)R}{1-V} & \text{if } V < \frac{1}{r_1} \\ 0 & \text{if } V \geq \frac{1}{r_1} \end{cases}$$

Clearly, when r_1 increases, r_2 will decrease. And it is also a decreasing function in V ,

since $\frac{\partial r_2}{\partial V} = -\frac{(r_1-1)R}{(1-V)^2} < 0$ and $\frac{(1-r_1V)R}{1-V} > 0$ when $V < \frac{1}{r_1}$.

Note that r_2 reaches its largest value, $r_2^* = \frac{(1-\lambda r_1)R}{1-\lambda}$ when V takes its lowest value of λ . And r_2 will be equal to r_1 when V takes the value of $\hat{V} = \frac{R-r_1}{R-1} \frac{1}{r_1} < \frac{1}{r_1}$. Obviously, \hat{V} is not an equilibrium point: if the patient agent believes actual V will be a bit smaller than \hat{V} , and she believes that all other patient agents have the same beliefs, then she will not demand early withdrawal since waiting until the last period generates a higher return. Other patient depositors will do the same, and V will equal λ , the good equilibrium is achieved. But if the patient agent believes actual V will be a bit larger than \hat{V} , and she believes that all other patient agents have the same beliefs, then she will demand early withdrawal since the return of waiting is already smaller than r_1 . And the more a patient depositor believes that V will be large, the stronger her incentive will be to demand early withdrawal. Other patient depositors will do the same, which reinforces the belief. This strategic complementarity in demanding early withdrawal will lead V to reach its maximum value, 1, and the bad equilibrium is achieved.

Note that when the return is revealed to be R_L , no one is willing to wait until $t = 2$, and the optimal choice is to liquidate the long-run investment, and each agent has the probability of $1/r_1$ to get the short-run payment r_1 .

The banks will set r_1 through the following optimization problem

$$\max (1-q)(1-s)[\lambda u(r_1) + (1-\lambda)u(r_2)] + [(1-q)s + q] \frac{1}{r_1} u(r_1) \quad (\text{B.3.6})$$

subject to the resource constraint

$$r_2 = \frac{(1-\lambda r_1)R_H}{1-\lambda} \quad (\text{B.3.7})$$

Table B.1: Consumption Patterns

Models	First-order condition	Resource constraint
Autarky	$c_m = 1$	$c_p = \begin{cases} R_H & \text{with prob. of } (1-q) \\ 1 & \text{with prob. of } q \end{cases}$
Optimal risk sharing	$u'(r_1^*) > R_H u'(r_2^*)$	$c_{p2}^* = \frac{(1-\lambda c_m)R_H}{1-\lambda}$
Deposit contract without insurance	$u'(r_1^*) > R_H u'(r_2^*)$	$r_2^* = \frac{(1-\lambda r_1^*)R_H}{1-\lambda}$
Deposit contract with insurance	$u'(r_1^{**}) = R_H u'(r_2^{**})$	$r_2^{**} = \frac{R_H[\frac{1}{1-q} - (\lambda + \frac{q}{1-q})r_1^{**}]}{1-\lambda}$

Table B.2: Welfare

Models	Welfare
Autarky	$W'_{autarky} = (1-q)[\lambda u(1) + (1-\lambda)u(R_H)] + qu(1)$
Optimal risk sharing	$W'_{optimal} = (1-q)[\lambda u(c_m^*) + (1-\lambda)u(c_{p2}^*)] + qu(1)$
Deposit contract without insurance	$W_{TSNI} = (1-q)(1-s)[\lambda u(r_1^*) + (1-\lambda)u(r_2^*)] + [(1-q)s + q]\frac{1}{r_1^*}u(r_1^*)$
Deposit contract with insurance	$W_{TSWI} = (1-q)[\lambda u(r_1^{**}) + (1-\lambda)u(r_2^{**})] + qu(r_1)$

which gives us

$$u'(r_1^*) - R_H u'(r_2^*) = \frac{(1-q)s + q}{\lambda(1-q)(1-s)} \frac{1}{(r_1^*)^2} [u(r_1^*) - r_1^* u'(r_1^*)] > 0 \quad (\text{B.3.8})$$

where

$$r_2^* = \frac{(1-\lambda r_1^*)R_H}{1-\lambda} \quad (\text{B.3.9})$$

B.3.3 Deposit Insurance

Now in this section, the banks will set up a deposit insurance fund and each puts in a . The sequential service constraint still applies, and the deposit insurance works the following way:

(1) If $V \leq \frac{1-a}{r_1}$, i.e., the banks can fulfill all withdrawal demand by liquidating assets, the deposit insurance fund will not intervene;

(2) If $V > \frac{1-a}{r_1}$, i.e., some early withdrawal demand can not be satisfied even when

banks already liquidate all assets, then the deposit insurance fund will ensure that those depositors with $f_i \leq \frac{1}{r_1}$ get r_1 , the deposit contract return in period $T = 1$, and that the rest of the depositors get no more than δ , the insurance coverage limit. Note that if $\delta \leq r_1$, then this limit works the same as a guarantee. If $\delta > r_1$, then it implies full insurance, and the actual insurance payout will be r_1 .

When deposit insurance is put in place, we have

$$r_2 = \begin{cases} \frac{(1-a-r_1V)R}{1-V} & \text{if } V < \frac{1-a}{r_1} \\ 0 & \text{if } V \geq \frac{1-a}{r_1} \end{cases}$$

Again, r_2 decreases in V , and will reach its largest value, $r_{2DI}^* = \frac{(1-a-\lambda r_1)R}{1-\lambda}$ when V takes its lowest value of λ . And r_2 will be equal to r_1 when V takes the value of $\hat{V}_{DI} = \frac{(1-a)R-r_1}{R-1} \frac{1}{r_1}$, which is smaller than $\frac{1}{r_1}$ and also smaller than \hat{V} .

Obviously, \hat{V}_{DI} is not an equilibrium point: if the patient agent believes actual V will be a bit smaller than \hat{V}_{DI} , and she believes that all other patient agents have the same beliefs, then she will not demand early withdrawal since waiting until the last period generates a higher return. Other patient depositors will do the same, and V will equal λ , the good equilibrium is achieved. But if the patient agent believes actual V will be a bit larger than \hat{V}_{DI} , and she believes that all other patient agents have the same beliefs, then she will demand early withdrawal since the return of waiting is already smaller than r_1 . And the more a patient depositor believes that V will be large, the stronger her incentive will be to demand early withdrawal, i.e., strategic complementarity. Other patient depositors will do the same, which reinforces the belief.

When V increases to the value of $\frac{1-a}{r_1}$, the deposit insurance will be triggered. Patient depositors still get zero if they choose to wait, but if they demand early withdrawal, there is a chance of $\frac{1}{r_1}$ of getting r_1 , and otherwise get δ , compared to the no insurance case where they get r_1 with probability of $\frac{1}{r_1}$, and otherwise get nothing.

From the above analysis, we see that with the deposit insurance, the threshold value of V which equate r_1 and r_2 becomes smaller, implying a potential higher probability of panic bank runs since the range of “run” beliefs is larger. The advantage of the insurance, however, is that it increases the welfare when bank run occurs.

The following section gives the mathematical representation of these two counteracting effects, therefore allowing us to analyze the welfare impact of deposit insurance and to find out the optimal level of the insurance coverage limit.

When bank run happens, the bank liquidates all its assets ($1 - a$ when there is no liquidation cost), and the deposit insurance fund will pay the gap between the liquidation value and the promised value, which is $\frac{1}{r_1}r_1 + (1 - \frac{1}{r_1})\delta - (1 - a) = (1 - \frac{1}{r_1})\delta + a$.

The actuarially fair premium ⁹ is:

$$a - \alpha[(1 - \frac{1}{r_1})\delta + a] = 0, \quad a = \frac{\alpha}{1 - \alpha}(1 - \frac{1}{r_1})\delta \quad (\text{B.3.10})$$

We argue that there will be no sunspot bank runs, and will check this proposition afterwards.

With the deposit insurance, the optimization problem of the bank is

$$\max (1 - q)[\lambda u(r_1) + (1 - \lambda)u(r_2)] + q[\frac{1}{r_1}u(r_1) + (1 - \frac{1}{r_1})u(\delta)] \quad (\text{B.3.11})$$

subject to

$$r_2 = \frac{R_H(1 - a - \lambda r_1)}{1 - \lambda} \quad \text{and} \quad a = \frac{q}{1 - q}(1 - \frac{1}{r_1})\delta \quad (\text{B.3.12})$$

⁹Note that this calculation method works only if there are only idiosyncratic bank runs, which are not realistic. We'll explore other premium calculation methods later, for instance, the worst-scenario premium, which means the premium collected must be no less than the liquidity needed in times of crisis. This method also assumes away credit line or bailout of the fund by central banks or treasuries. We consider the actuarially fair case first for simplicity.

Taking first-order derivatives, we'll get the following two conditions:

$$u'(r_1) - R_H u'(r_2) = \frac{q}{1-q} \frac{1}{\lambda r_1^2} [u(r_1) - r_1 u'(r_1) - u(\delta) + \delta R_H u'(r_2)] \quad (\text{B.3.13})$$

$$u'(\delta) = R_H u'(r_2) \quad (\text{B.3.14})$$

Therefore, we have $u'(r_1) - u'(\delta) = \frac{q}{1-q} \frac{1}{\lambda r_1^2} [u(r_1) - r_1 u'(r_1) - u(\delta) + \delta u'(\delta)]$. Since $u(x) - x u'(x)$ is increasing in x for the concave function u , the RHS is positive when $r_1 > \delta$. But the LHS will become negative since $u'(x)$ is decreasing in x . The equation will hold only when $\delta = r_1$, which implies full insurance¹⁰.

The optimal contract is given by

$$u'(r_1^{**}) = R_H u'(r_2^{**}) \quad (\text{B.3.17})$$

and

$$r_2^{**} = \frac{R_H \left[\frac{1}{1-q} - \left(\lambda + \frac{q}{1-q} \right) r_1^{**} \right]}{1 - \lambda} \quad (\text{B.3.18})$$

¹⁰Interestingly, if we change the implementation of the sequential service constraint and the design of the deposit insurance a little bit, we'll still get the same conclusion that the optimal level of insurance coverage limit is equal to the interest rate in period $T = 1$ promised in the demand deposit contract, i.e., $\delta = r_1$. The change is this: instead of serving depositors one by one in the line immediately, each depositor will get a number indicating her position in the line. Then, after all depositors who want to withdraw money in period $T = 1$ show up, the bank will count the number of depositors demanding early withdrawal. If that number is smaller than $\frac{1-a}{r_1}$, then deposit insurance will not be triggered, and the banks will liquidate their assets in order to meet the liquidity demand. Otherwise, the deposit insurance fund will intervene, and pay all depositors the same amount $\min\{\delta, r_1\}$.

In this case, the insurance payout and therefore the premium charged on banks are totally different. The actuarially fair premium will be $a = \alpha[\delta - (1-a)] = \frac{\alpha}{1-\alpha}(\delta - 1)$.

Then the banks' optimization problem becomes

$$\max (1-q)[\lambda u(r_1) + (1-\lambda)u(r_2)] + q u(\delta) \quad (\text{B.3.15})$$

subject to

$$r_2 = \frac{R_H(1-a-\lambda r_1)}{1-\lambda} \quad \text{and} \quad a = \frac{q}{1-q}(\delta - 1) \quad (\text{B.3.16})$$

Taking first-order derivatives, we'll get the following two conditions $u'(r_1) = R_H u'(r_2)$ and $u'(\delta) = R_H u'(r_2)$, which gives us $\delta = r_1$, i.e., full insurance, the same conclusion we drawn from the "withdraw as you go" case. And the values of r_1 , r_2 , a and δ will also be the same.

Going back to check whether the deposit insurance can indeed eliminate sunspot bank runs. Under deposit insurance, when the state is revealed to be R_H , if the patient agent decides to demand early withdrawal, she will get $u(r_1^{**})$ no matter whether there are bank runs or not, and if she decides to wait, she will get $u(r_2^{**})$. Equation (B.3.17) implies r_2^{**} is larger than r_1^{**} , therefore in the solvent state, not running on banks is the dominant strategy for patient depositors under deposit insurance. We have proved that sunspot runs will be eliminated and therefore the optimization problem is correct.

And of course we need to check whether the welfare under deposit insurance is higher than autarky to make sure that banks with deposit insurance are desirable in the very beginning.

Compared to the benchmark case, here the demand deposit contract accompanied with deposit insurance does *not* achieve optimal risk sharing. This is because the positive probability of bank runs incurs a premium charged on the banks, which reduces their profits and therefore the interest rate paid to depositors at period 1. This results in a more than optimal probability of bank runs, and therefore the welfare is reduced.

We assume that the utility function takes the following CARA form ¹¹

$$u(c) = 1 - e^{-\gamma c} \tag{B.3.19}$$

Plug the utility function into first-order equations and solve for the values, and we'll get numerical solutions for the optimizing problems in sections above. These values are computed in MATLAB, and the codes are listed in Appendix B.

The parameters take the following values in the numerical example:

The equilibrium values of variables are summarized in the table below.

When there is uncertainty about the true states of the world, the fact that bank run is bound to occur in the low state forces the banks to lower their non-contingent short-

¹¹Note that we require $u(0) = 0$ in our model. That's why the common form of utility function $u(c) = c^{1-\gamma}/(1-\gamma)$ doesn't apply here.

Table B.3: Parameter Values

Two-State Model		
Parameters	Value	
q	0.05	probability of low state
s	0.05	probability of sunspot
λ	0.5	proportion of impatient agents
γ	2	risk aversion
R_H	3	return to investment in high state
N	500	number of grids

Table B.4: Demand Deposit Contract

Two-State Model				
	Autarky	Optimal risk sharing	Demand Deposit Contract	
			Without Insurance	With Insurance
r_1	1	1.363	1.248	1.336
r_2	3	1.912	2.255	1.885
<i>Welfare</i>	0.928	0.952	0.932	0.953

run payment to the depositors, so that when bank run does occur, the welfare loss due to nothing to get for some agents can be reduced. This introduces a new type of risk sharing, namely **risk sharing across states**, into the classical Diamond-Dybvig model, which only has risk sharing among agents. Here, optimal contract is determined by a trade-off between risk-sharing across agents during the good state and reducing welfare loss during the bad state; in other words, the risk-sharing across late and early consumers is reduced to make room for risk-sharing across good and bad economic states in order to achieve the optimal allocation.

When the low state is realized, the chance is higher for each agent to get the short-run payment. This implies a risk sharing between different states of the world, in addition to the risk sharing between different types of agents. However, the across-state risk sharing comes at a cost. When high state is realized, there is less risk sharing between patient and impatient agents since $r_1 < r_{1H}$.

Also note that the expected welfare generating from the bank's deposit contract is higher than that of autarky as long as the probabilities s and q are small enough. The advantages brought by the risk-sharing function of the demand deposit contract outweighs the disadvantage resulting from bank runs.

B.3.4 Continuous-State Extension

Motivation: in the two-state model, the probability of insolvent state is an exogenous number, independent of demand-deposit contract interest rate. In the following section, we'll study the case where there is a continuum of states and where the probability of insolvent state and the magnitude of deposit interest rate are jointly determined.

There is a continuum of potential states and each corresponds to a risky return of R_i . We assume that the expected value of long-run payment, $E(r_2) = E(R)(1 - \lambda r_1)/(1 - \lambda)$ is greater than the short-run payment, therefore banks are still desirable *ex ante*.

The true state of the world will be revealed at $t = 1$, and each agent will act according to his or her type and the true state. Early consumers will always withdraw at $t = 1$, and late consumers will do so as well if R is too low.

For the sections below, we assume R follows uniform distribution ¹².

In Autarky state, if R is revealed to be smaller than one, both types of agents will liquidate their investment. If R is larger than one, only impatient agents will liquidate investment, and patient agents will wait until period 2 and consume R units of goods.

The threshold return \hat{R} is equal to one. The probability that both types of agents will demand early liquidation is

$$\hat{p} = Prob(R \leq 1) = \frac{1 - \underline{R}}{\bar{R} - \underline{R}} \quad (\text{B.3.20})$$

which is a constant.

¹²We will also explore the case where the distribution of investment return is a function of banker's managing efforts to study the effect of moral hazard when deposit insurance is put in place.

The social welfare is

$$\hat{W}_{autarky} = (1 - \hat{p})[\lambda u(1) + (1 - \lambda)E_{[1, \bar{R}]}u(R)] + \hat{p}u(1) \quad (\text{B.3.21})$$

where

$$E_{[1, \bar{R}]}u(R) = \int_1^{\bar{R}} \frac{1}{\bar{R} - 1} u(R_i) dR \quad (\text{B.3.22})$$

The optimal risk sharing question is

$$\max(1 - \hat{p})[\lambda u(c_m) + (1 - \lambda)E_{[1, \bar{R}]}u(c_{pi})] + \hat{p}u(1) \quad (\text{B.3.23})$$

where

$$E_{[1, \bar{R}]}u(c_{pi}) = \int_1^{\bar{R}} \frac{1}{\bar{R} - 1} u(c_{pi}) dR \quad (\text{B.3.24})$$

subject to the resource constraint

$$\lambda c_m + \frac{(1 - \lambda)c_{pi}}{R_i} = 1 \quad (\text{B.3.25})$$

for every $R_i \in [R, \bar{R}]$.

The first order equation is

$$u'(c_m^*) = \frac{1}{\bar{R} - 1} \int_1^{\bar{R}} u'(c_{pi}^*) R_i dR \quad (\text{B.3.26})$$

where

$$c_{pi}^* = \frac{(1 - \lambda c_m^*) R_i}{1 - \lambda} \quad (\text{B.3.27})$$

The welfare is

$$\hat{W}_{optimal} = (1 - \hat{p})[\lambda u(c_m^*) + (1 - \lambda)Eu(c_p)] + \hat{p}u(1) \quad (\text{B.3.28})$$

With the demand deposit contract, the threshold return \hat{R} , makes the late consumer indifferent between withdrawing money at $t = 1$ and at $t = 2$, which implies

$$\frac{1}{r_1}u(r_1) = u(r_2) = u\left(\frac{\hat{R}(1 - \lambda r_1)}{1 - \lambda}\right) \quad (\text{B.3.29})$$

therefore

$$\hat{R} = \hat{R}(r_1) = \frac{(1 - \lambda)u^{-1}\left(\frac{1}{r_1}u(r_1)\right)}{1 - \lambda r_1} \quad (\text{B.3.30})$$

When $R > \hat{R}$, the economic fundamentals are good enough, and there will only be panic-based bank runs; namely sun-spot equilibrium. In DD model, they assume an exogenous probability of this kind of bank runs. We can assume the good equilibrium happens all the time first for simplicity.

When $R \leq \hat{R}$, economic fundamentals are so bad that both late and early consumers will run on the banks. In other words, there is bound to be fundamental-driven bank runs.

When there is no bank run, only early consumers demand withdrawal at $t = 1$ and each get r_1 . Late consumers will wait until $t = 2$ and get whatever is left when the return of the risky project is realized, $r_2 = \frac{1 - \lambda r_1}{1 - \lambda} R$.

When there is a bank run and no deposit insurance, due to the sequential service constraint, each agent has the probability of $1/r_1$ to get the short-run payment r_1 , and get zero otherwise.

Same as the two-state model, we assume R follows a uniform distribution over $[\underline{R}, \bar{R}]$. Therefore, the probability with which $R < \hat{R}$ is

$$\hat{p} = \text{prob}(R < \hat{R}) = \frac{1}{\bar{R} - \underline{R}} \hat{R} - \frac{\underline{R}}{\bar{R} - \underline{R}} \quad (\text{B.3.31})$$

The optimizing problem is

$$\max (1 - \hat{p})(1 - s)[\lambda u(r_1) + (1 - \lambda)E_{[\hat{R}, \bar{R}]}u(r_{2i})] + [(1 - \hat{p})s + \hat{p}]\frac{1}{r_1}u(r_1) \quad (\text{B.3.32})$$

where

$$E_{[\hat{R}, \bar{R}]}u(r_{2i}) = \int_{\hat{R}}^{\bar{R}} \frac{1}{\bar{R} - \hat{R}} u(r_{2i}) dR \quad (\text{B.3.33})$$

and

$$r_{2i} = \frac{(1 - \lambda r_1)R_i}{1 - \lambda} \quad (\text{B.3.34})$$

for R_i greater than \hat{R} .

The first-order condition is

$$\begin{aligned} \lambda[u'(r_1) - \frac{1}{\bar{R} - \hat{R}} \int_{\hat{R}}^{\bar{R}} u'(r_{2i}) R_i dR] &= \frac{1}{1 - \hat{p}} \frac{d\hat{p}}{dr_1} [\lambda u(r_1) + (1 - \lambda)E_{[\hat{R}, \bar{R}]}u(r_{2i}) - \frac{1}{r_1}u(r_1)] \\ &+ \frac{s + \frac{\hat{p}}{1 - \hat{p}}}{1 - s} \frac{1}{(r_1)^2} [u(r_1) - r_1 u'(r_1)] \\ &+ \frac{d\hat{R}}{dr_1} \frac{1 - \lambda}{\bar{R} - \hat{R}} \frac{1}{r_1} u(r_1) \end{aligned} \quad (\text{B.3.35})$$

There is a deposit insurance which charges a premium of a at $t = 0$ and ensures each depositor to get δ when there is a bank run.

Same as in the two-state model, we suppose the premium is charged on the bank side. There is a new threshold return \tilde{R} , which makes the late consumer indifferent between withdrawing money at $t = 1$ and at $t = 2$ with deposit insurance,

$$u(\delta) = u(\tilde{r}_2) = u\left(\frac{\tilde{R}(1 - \lambda r_1)}{1 - \lambda}\right) \quad (\text{B.3.36})$$

which implies

$$\delta = \tilde{r}_2 = \frac{\tilde{R}(1 - \lambda r_1)}{1 - \lambda} \quad (\text{B.3.37})$$

therefore

$$\tilde{R} = \tilde{R}(\delta, r_1) = \frac{(1 - \lambda)\delta}{1 - \lambda r_1} \quad (\text{B.3.38})$$

The optimization problem under deposit insurance is

$$\max (1 - \tilde{p})[\lambda u(r_1) + (1 - \lambda)E_{[\tilde{R}, \bar{R}]}u(r_{2i})] + \tilde{p}u(\delta) \quad (\text{B.3.39})$$

where

$$E_{[\tilde{R}, \bar{R}]}u(r_{2i}) = \int_{\tilde{R}}^{\bar{R}} \frac{1}{\bar{R} - \tilde{R}} u(r_{2i}) dR \quad (\text{B.3.40})$$

subject to

$$r_{2i} = \frac{(1 - a - \lambda r_1)R_i}{1 - \lambda} \quad (\text{B.3.41})$$

where

$$a = \frac{\tilde{p}}{1 - \tilde{p}}(\delta - 1) \quad (\text{B.3.42})$$

First-order conditions with respect to r_1 and δ are, respectively

$$\begin{aligned} & \lambda[u'(r_1) - \frac{1}{\bar{R} - \tilde{R}} \int_{\tilde{R}}^{\bar{R}} u'(r_{2i}) R_i dR] \\ &= \frac{\partial \tilde{R}}{\partial r_1} \frac{1 - \lambda}{\bar{R} - \tilde{R}} u(\tilde{r}_2) + \frac{1}{1 - \tilde{p}} \frac{\partial \tilde{p}}{\partial r_1} [\lambda u(r_1) + (1 - \lambda)E_{[\tilde{R}, \bar{R}]}u(r_{2i}) - u(\delta)] \end{aligned} \quad (\text{B.3.43})$$

and

$$\begin{aligned} \tilde{p}u'(\delta) &= (1 - \tilde{p})(1 - \lambda) \frac{1}{\bar{R} - \tilde{R}} \left[\frac{\partial \tilde{R}}{\partial \delta} u(\tilde{r}_2) - \int_{\tilde{R}}^{\bar{R}} u'(r_{2i}) \frac{\partial r_{2i}}{\partial \delta} dR \right] \\ &+ \frac{\partial \tilde{p}}{\partial \delta} [\lambda u(r_1) + (1 - \lambda)E_{[\tilde{R}, \bar{R}]}u(r_{2i}) - u(\delta)] \end{aligned} \quad (\text{B.3.44})$$

Plug the utility function into first-order equations and solve for the values, and we'll get numerical solutions for the optimizing problems in sections above.

Table B.5: Parameter Values

Continuous-State Model		
Parameters	Value	
s	0.05	probability of sunspot
λ	0.5	proportion of impatient agents
γ	2	risk aversion
\bar{R}	6	maximum return to investment
\underline{R}	0	minimum return to investment

Table B.6: Demand Deposit Contract

Continuous-State Model				
	Autarky	Optimal risk sharing	Demand Deposit Contract	
			Without Insurance	With Insurance
r_1	1	1.215	1.109	0.219
$E(r_2)$	3.5	2.747	3.08	5.975
<i>Welfare</i>	0.915	0.93	0.909	0.704

Appendix C

Appendix of Chapter 3

C.1 First-best allocation

A benevolent social planner with perfect information will maximize the expected welfare:

$$\max E[\lambda U(c_{m1}) + (1 - \lambda)\beta U(c_{p1} + c_{p2})]$$

subject to the resource constraint

$$(1 - \lambda)c_{p2} = \left[1 - \frac{\lambda c_{m1} + (1 - \lambda)c_{p1}}{y_i}\right] Y_i$$

Plug $c_{p2} = \frac{1 - \frac{\lambda c_{m1}}{y_i}}{1 - \lambda} Y_i$ to the utility function

Safe projects:

$$\max \lambda U(c_{m1}) + (1 - \lambda)\beta U\left(\frac{1 - \frac{\lambda c_{m1}}{y_s}}{1 - \lambda} Y_s\right)$$

first-order condition

$$\begin{aligned} 0 &= \lambda U'(c_{m1}) + (1 - \lambda)\beta U'(c_{p2}) \frac{-\lambda}{1 - \lambda} \frac{Y_s}{y_s} \\ &= U'(c_{m1}) - \beta U'(c_{p2}) \frac{Y_s}{y_s} \end{aligned}$$

Hence, the two equations solving the optimization problem are

$$U'(c_{m1}^s) = \beta \frac{Y_s}{y_s} U'(c_{p2}^s)$$

$$c_{p2}^s = \frac{1 - \frac{\lambda c_{m1}^s}{y_s}}{1 - \lambda} Y_s$$

Risky projects:

$$\max \lambda U(c_{m1}) + (1 - \lambda)\beta q U\left(\frac{1 - \frac{\lambda c_{m1}}{y_r}}{1 - \lambda} Y_r\right)$$

first-order condition

$$\begin{aligned} 0 &= \lambda U'(c_{m1}) + (1 - \lambda) \beta q U'(c_{p2}) \frac{-\lambda Y_r}{1 - \lambda y_r} \\ &= U'(c_{m1}) - \beta q U'(c_{p2}) \frac{Y_r}{y_r} \end{aligned}$$

Hence, the two equations solving the optimization problem are

$$U'(c_{m1}^r) = \beta \frac{q Y_r}{y_r} U'(c_{p2}^r)$$

$$c_{p2}^r = \frac{1 - \frac{\lambda c_{m1}^r}{y_r}}{1 - \lambda} q Y_r$$

If $c_{m1}^r = c_{m1}^s$, comparing the resource constraints, then $c_{p2}^r < c_{p2}^s$ since $y_r < y_s$ and $q Y_r < Y_s$.

If $c_{p2}^r = c_{p2}^s$, comparing the first-order conditions, then $c_{m1}^r < c_{m1}^s$ since $U'(c_{m1}^r) > U'(c_{m1}^s)$ if $\frac{q Y_r}{y_r} > \frac{Y_s}{y_s}$, namely if $y_s > \frac{y_r}{q Y_r} Y_s$.