ABSTRACT

Title of Document: ESSAYS ON ASSET PRICING AND FINANCIAL STABILITY

Jeong Min Lee, Doctor of Philosophy, 2014

Directed by: Professor Albert S. Kyle
Professor Mark Loewenstein
Department of Finance
Robert H. Smith School of Business

My two-essay dissertation revolves around understanding the financial crisis of 2008. First I focus on the repo market, a major funding source of the shadow banking system, and show the repo market can create and amplify the fragility of the system. Then I investigate a broader economy with heterogeneous agents and demonstrate how the dynamics of equilibrium asset prices and wealth distributions are determined.

In Essay 1, I develop a dynamic model of collateral circulation in a repo market, where a continuum of institutions borrow from and lend to one another against illiquid collateral. The model emphasizes an important tradeoff. On one hand, easier collateral circulation makes repos liquid and increases steady state investment through several multiplier effects, improving economic efficiency. On the other hand, it can harm financial stability because less capital is sitting on the sidelines waiting for investment opportunities. This fragility is further exacerbated by the endogenous repo spread through a positive feedback loop, and can result in an inefficient repo run. The model is relevant for understanding the repo markets

during the financial crisis of 2008.

In Essay 2, I study the dynamics of the wealth distribution and asset prices in a general equilibrium model. Agents face heterogeneous portfolio constraints that limit the shares of risky investments relative to wealth. The setup is motivated by empirical evidence that many households do not participate in the stock market and portfolio shares are heterogeneous and persistent conditional on stock market participation. There are two main results. First, one state variable can summarize the wealth distribution regardless of the number of types of agents. Second, when the economy is bad, it becomes more sensitive to additional negative shocks, meaning that not only magnitudes of the shocks but also their frequency matters.

ESSAYS ON ASSET PRICING AND FINANCIAL STABILITY

by

Jeong Min Lee

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:
Professor Albert S. Kyle, Co-Chair
Professor Mark Loewenstein, Co-Chair
Professor Richmond Mathews
Professor Haluk Unal
Professor John Shea

Dedication

I dedicate this dissertation to my parents, Jongtae Lee and Keunae Kwak.

Acknowledgments

This dissertation would not have been possible without the help and support of many people. First and foremost I want to thank my advisors, Pete Kyle and Mark Loewenstein, for their constant guidance, advice, support and encouragement through the six years of my PhD study. They have always been great mentors, inspiring scholars, and the best role models. I also thank my committee members, Rich Mathews, Haluk Unal and John Shea, for their insights and support throughout the process. I thank alll the faculty members and PhD students in the Finance Department for their help and kindness. Last but not least I am grateful to my husband, Leland Crane, whose love has been essential through this process.

Contents

List of Figures

Chapter I: Collateral Circulation and Repo Spreads

1 Introduction

The financial crisis of 2008 has highlighted the fact that procyclical leverage can be devastating in downturns. Assets and liabilities of financial institutions, often reflecting collateralized lending mechanisms such as repos, expanded and contracted with market fluctuations. Scholars have suggested that this leverage cycle might arise due to price impact, fluctuating risk, shifts in beliefs of the marginal buyer, changes in information asymmetry, or Knightian uncertainty.[1]

Less emphasized is the fact that financial institutions circulate (repledge or rehypothecate) collateral. One institution may be lending to another institution against some collateral, while borrowing against the same collateral pledged by his borrower. This aspect of the financial system has been understudied. In particular, most models in the literature feature two distinct groups, one that lends and one that borrows, and focus only on borrowers' financing constraints.

During the crisis, collateral circulation contracted more than leverage itself. Figure I.1 plots, for the major US broker-dealers, the total book liabilities and total value of collateral available for circulation over the last ten years. Between 2003 and

[1] See Kiyotaki and Moore (1997b), Cifuentes, Ferrucci, and Shin (2005), Shin (2010), Adrian and Shin (2011), Gromb and Vayanos (2002), Brunnermeier and Pedersen (2009), Geanakoplos (2010), Dang, Gorton, and Holmstrom (2012), Gorton and Ordonez (2012), Caballero and Krishnamurthy (2008), Caballero and Simsek (2009), among others.

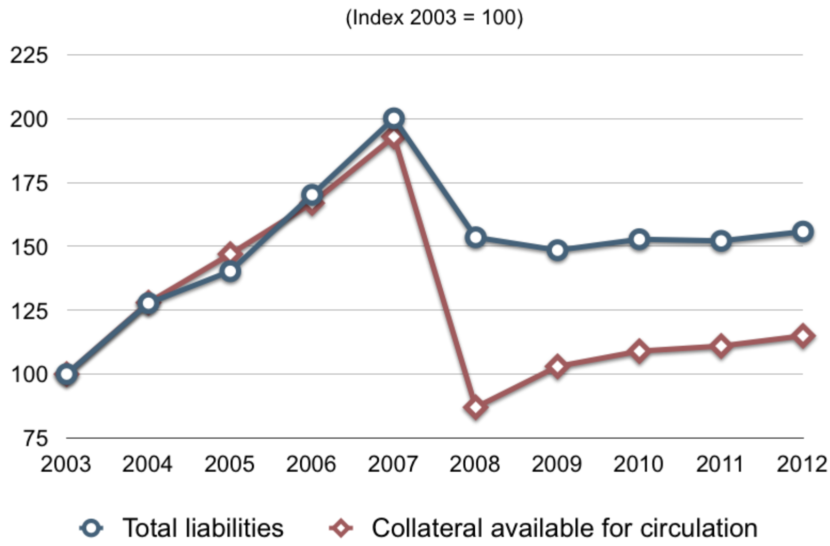


Figure I.1: The data is from 10-K reports of major US broker-dealers: Morgan Stanley, Merrill Lynch/Bank of America, JP Morgan/Bear Stearns, Lehman Brothers, and Citigroup.

2007, liabilities and collateral circulation both doubled in value. In 2008, a quarter of liabilities evaporated, as did *half* the value of collateral available for circulation. In absolute terms, the decline in the value of collateral for circulation was about 2.5 trillion dollars.[2]

My goal in this paper is to understand the economic importance of collateral circulation. Does circulating collateral only mean the same collateral is counted multiple times on the book, an accounting convention devoid of economic significance? Does allowing lenders and borrowers to be the same entities dampen or amplify the leverage cycle? What are the implications of lenders, as well as borrowers, facing financing constraints? What happens when collateral circulation suddenly declines?

To address these questions I build a dynamic model of collateral circulation. I find

[2] Singh (2011) calculates the length of collateral chains using hedge fund data and reports collateral was circulated on average 3 times in 2007 and 2.4 times in 2010. Singh and Aitken (2009, 2010) document the sudden shortening of collateral chains in the US and Europe. See also Monnet (2011).

that collateral circulation poses a tradeoff between economic efficiency and financial stability. While easier collateral circulation makes repos liquid and increases steady state investment through several multiplier effects, it can make the financial system more fragile because less capital is left sitting on the sidelines waiting for opportunities to arrive. This fragility is exacerbated by a positive feedback loop between the endogenous repo spread and fire-sale discounts. As a result, a sudden contraction of collateral circulation can result in an inefficient repo run.

In the model, there is a continuum of ex-ante symmetric financial institutions, the total mass of which is fixed at unity. Each institution is endowed with the same amount of liquid capital (cash) and cannot raise outside equity. These institutions invest in one class of assets that are *illiquid*. They can scale up their investments by pledging the assets as collateral in order to borrow from other institutions. Thus all investments are financed within the system of financial institutions, which has limited liquid capital. This setup is motivated by the recent finding of Krishnamurthy, Nagel, and Orlov (2012) that a large share of complex and illiquid securities, such as subprime mortgage-backed securities, are financed by sophisticated financial institutions such as broker-dealers, investment banks, hedge funds, or off-shore funds (or shadow banks).

Financial institutions receive random arrivals of investment opportunities as in Kiyotaki and Moore (2001). An investment opportunity allows the institution to purchase risky assets. The opportunity is specific to the institution, reflecting the institution's specialized skills. Assets mature at a random time and have to be sold at a discount in the secondary market if liquidated before maturity. These investments can be interpreted as originating and seasoning a pool of mortgages.

Due to the random arrival of investment opportunities, at any given point in time some financial institutions have opportunities while others do not. The *repo market* allows those with opportunities to borrow from those without. Repos are collateralized lending mechanisms where the haircut (i.e., the percentage difference between the principal and the collateral value) is set to make the loan safe. In particular, I assume the haircut is set sufficiently high so that lenders can recover the principal by liquidating the collateral in the event of default by the borrower. Financial institutions with opportunities optimally choose whether and how much to borrow from the repo market. Financial institutions without opportunities decide whether to invest in cash, earning the riskless rate, or to invest in repos, earning the repo interest rate. The composition of the financial system (i.e., the masses of asset, cash, and repo investors) is also endogenous.

When a repo investor receives an investment opportunity, he seeks to rehypothecate (i.e., circulate) collateral pledged by his borrower in order to finance his investment. I assume that, with some probability q, the lender is able to circulate the collateral quickly enough to finance the investment opportunity; with probability $1-q$, the lender is unable to circulate the collateral, and the investment opportunity is lost. The value of q thus measures the liquidity or "moneyness" of repo collateral, and $1-q$ measures frictions associated with repos and not with cash. Repos are equivalent to cash if $q=1$.

Repo liquidity, q, captures three aspects of real-world collateral circulation: (a) borrowers who worry about getting their collateral back may not allow lenders to circulate it; (b) lenders may not circulate collateral for precautionary motives; and (c) the ability of one institution to quickly circulate collateral is limited by its connections to other institutions.

Taking the value of q as constant, I first study how the steady state of a dynamic competitive economy depends on a given value of q. I then study how completely unanticipated shocks, such as changes in q, affect the transition path from one steady state to another.[3]

Two main results emerge in a stationary equilibrium. First, the endogenous spread between the repo rate and the riskless rate is generally positive, although repos are free of credit risk. The positive repo spread compensates lenders for the potential illiquidity of collateral when $q < 1$. The spread decreases in repo liquidity, and increases in the arrival rate and the profitability of investment opportunities. This positive spread is consistent with empirical evidence documented during the recent financial crisis.[4] Moreover, this result stands in contrast to the literature following Duffie (1996) on repo specialness, which implies negative spreads. Since the financial system has limited cash, the cost of financing illiquid investments can be higher than that of liquid investments, as captured by the repo spread.

Second, increased repo liquidity raises aggregate investment through three multiplier effects. When repos are highly liquid, (a) each institution with an opportunity can take higher leverage and scale up their investment; (b) it is less likely to default despite high leverage because the repo spread is sufficiently low; and (c) more institutions of the financial system can participate in risky investment. Therefore, repo liquidity creates real multiplier effects through the endogenous choices of financial institutions.

After analyzing the steady state, I introduce unanticipated, permanent shocks and

[3] Ongoing work suggests the results are largely similar when the shocks are anticipated and sufficiently infrequent.

[4] See Gorton and Metrick (2011), Hordahl and King (2008), and Smith (2012).

study the resulting deterministic transition paths from one steady state to another. I focus on negative shocks to repo liquidity and the rate of asset maturity (i.e., economic slowdowns). I find two types of transition paths: a *stable path* and a *repo run*. With a relatively small shock the secondary market absorbs the increased supply of liquidated assets and the haircut remains constant, generating a stable path. In contrast, a larger shock causes a repo run where collateral becomes worthless, the haircut shoots up to unity, and all leveraged investors go bankrupt. A positive feedback effect between the repo spread and fire-sale discounts is responsible for a repo run. When a shock causes an excess supply of liquidated collateral, the secondary market price of these assets adjusts downward, resulting in fire-sale discounts and higher haircuts. To acquire cheap liquidated assets, financial institutions with opportunities need to borrow from other institutions in the system, who also understand the profit opportunities. The fire-sale discounts increase the repo spread further, as the opportunity cost of having to forego profitable investments is higher. The escalating repo spread pressures existing borrowers and discourages new borrowers from taking high leverage, leading to deeper discounts and higher haircuts. Thus, lenders and borrowers being the same entities can be destabilizing when the cost of financing is endogenous.

The amount of cash in the financial system determines whether a given shock triggers a repo run. When repos have little moneyness the financial system hoards large amounts of cash. The cash reserves, although "inefficient" when compared to an equilibrium with more repo moneyness, make the financial system more resilient. When there is a shock, the cash reserves help buffer the shock along the transition path to the new equilibrium. When reserves are insufficient, a shock results in fire sales and a repo run.

A repo run is constrained inefficient, in that a social planner with the same collateral circulation constraint can achieve Pareto improvement. The possibility of an inefficient repo run implies a role for a lender of last resort. To prevent a repo run, central banks can lend to solvent financial institutions against a variety of collateral at low rates, and with high haircuts to make loans safe. Such funding programs reward institutions for being solvent in crises, mitigating moral hazard problems.

My model sheds light on the divergence across repo markets during the financial crisis of 2008. There are two types of repos: tri-party and bilateral repos. Tri-party repos resemble traditional demand deposits without deposit insurance; and collateral does not circulate. In bilateral repos, as in my model, sophisticated financial institutions borrow from and lend to one another and circulate collateral. During 2008, bilateral repos contracted sharply, as documented by Gorton and Metrick (2010a,b, 2011, 2012). Both the haircut and the repo spread rose markedly. Certain classes of securities stopped being used as collateral entirely, and the repo spread rose from under 10 bp to over 200 bp on average, with a maximum around 700 bp (see I.10). On the other hand, Krishnamurthy, Nagel, and Orlov (2012) find tri-party repos remained largely stable throughout the crisis. Consistent with their findings, my model shows that bilateral repos can be fragile, with collateral circulation exacerbating the leverage cycle.

1.1 Related Literature

In the context of a repo market where collateral consists of Treasury securities, Duffie (1996) shows that the repo spread can be negative (or repo "special").[5] The negative spread compensates collateral owners (who are borrowers in my model) for lending scarce securities because short-sellers need to borrow the securities. In contrast to Duffie (1996), in my model liquid capital is scarce and the positive repo spread compensates cash lenders. I contribute to this literature by modeling circulating repo collateral and explaining positive repo spreads.

Monnet and Narajabad (2012) study the role of repos as opposed to asset sales. They find that repos coexist with asset sales when there are bilateral trading frictions and the uncertainty in the future value of collateral. Repos in my model are used to finance investment because a given asset is worth more to borrowers than to lenders, due to its inherent illiquidity. This role of repos as a funding source for financial institutions has been emphasized since the recent financial crisis. Martin, Skeie, and von Thadden (2011) focus on tri-party repos and show that repo runs may result from investors not being able to adjust the haircuts quickly. In my model, the terms of repos are updated continuously, yet repo runs may still occur. In addition, my model produces systemic failures of the financial sector rather than focusing on a single institution.

More broadly, my fragility result is related to the bank run literature following Diamond and Dybvig (1983). They show that banks' capital structures, with illiquid assets but liquid liabilities (demand deposits), are vulnerable to changes in credi-

[5]See also Duffie, Garleanu, and Pedersen (2002), Krishnamurthy (2002), Vayanos and Weill (2008) and Bottazzi, Luque, and Páscoa (2012).

tors' perceptions. Otherwise solvent banks may have to fail as each creditor fears that others might run first. My model differs from theirs in that all financial institutions are rational and the lenders understand fire-sale discounts, which actually makes the system more fragile. Relatedly, Brunnermeier and Pedersen (2009) show the fragility stemming from a positive feedback loop between the liquidation price and financing constraints. He and Xiong (2012a) and Acharya, Gale, and Yorulmazer (2011) emphasize short debt maturity as a source of fragility. The primary difference between these papers and mine is that I make the cost of financing endogenous. In my model this repo spread affects the amount of cash reserves in the system, and plays a critical role in generating fragility.

Collateral circulation in my model is related to traditional credit chains. Rochet and Tirole (1996) and Kiyotaki and Moore (1997a) study an economy where firms (or banks) borrow from and lend to one another. They show how credit chains can lead to systemic risk as a liquidity shock to one firm in the network spreads through chain reactions of default, i.e., domino effects.[6] In contrast, systemic risk of collateral circulation arises from the endogenous responses of both borrowers and lenders, rather than from chains of defaults. My paper is also related to the broader literature on interbank markets, where banks makes unsecured loans to one another.[7]

Finally, this paper is related to the literature on private money creation in monetary economics. With circulating collateral, the repo market effectively creates money. Townsend and Wallace (1987) motivate circulating collateral by location

[6]See also Acemoglu, Ozdaglar, and Tahbaz-Salehi (2013), Gofman (2011), and Zawadowski (2013).

[7]The fragility and inefficiency of these interbank markets have been studied by Bhattacharya and Gale (1987), Flannery (1996), Allen and Gale (2000), Repullo (2005), Acharya and Skeie (2011), and Acharya, Gromb, and Yorulmazer (2012), among others.

mismatch, while Kiyotaki and Moore (2001) use timing mismatch. The inefficiency of repo runs in my model is consistent with Friedman (1960), who strongly argues against the private creation of money. He argues that allowing private provision of circulating liabilities generates indeterminacy of equilibrium and excess volatility. Hayek (1976) and Fama (1980) oppose Friedman's position, arguing that the creation of money can be done in markets efficiently. My model implies that although collateral circulation makes the economy fragile and exposes it to the possibility of a repo run, it creates positive multiplier effects in normal conditions. Relatedly, Kiyotaki and Moore (2012) show fiat money can lubricate the economy when there are borrowing and resaleability constraints. They also show that the expected rate of return on money is lower than that of equity, which is also lower than the time discount rate. In their model, only money circulates and there is no debt. Stein (2012) study the link between monetary economics and financial stability where commercial banks create private money by issuing short-term debts. I contribute to this literature by showing how the shadow banking system can create private money through circulating collateral in repo markets.

I proceed by describing the model in the next section and solving the stationary equilibrium in Section 3. Section 4 studies transition paths and fragility. Section 5 discusses the inefficiency of a repo run and policy implications, while Section 6 serves as a conclusion.

2 Model

Time is continuous with an infinite horizon. There is a continuum of financial institutions (FIs) whose total mass is normalized to 1. All FIs are risk-neutral and have a time discount parameter of ρ, which is also the riskless rate. Each FI is endowed with K units of liquid capital (cash), a numeraire. FIs can either keep their cash in checking accounts, earning riskless rate ρ (whom I call **cash investors**), or invest in the repo market, earning repo rate $\rho+s$ with s being the spread between the repo rate and the riskless rate (whom I call **repo investors**). All repos mature at a random time that arrives according to a Poisson process with rate σ. Repo investors can switch to cash investors once their repos mature. FIs receive an investment opportunity at a random time that arrives according to a Poisson process with rate α. An opportunity is specific to the FI to which it arrives and reflects its special skills. If cash investors receive an opportunity, they instantly take it and become **(asset) investors**. However, if repo investors receive an opportunity before their repos mature, they should borrow cash against their borrower's collateral (i.e., circulating the collateral) to take the investment opportunity. A repo investor can do so only with exogenous probability $q \in [0,1]$, which measures the liquidity or moneyness of repo.

Investors can either buy new assets from the primary market at price B or buy liquidated assets from the secondary market at $p+\phi$, where ϕ is the exogenous per-unit transaction cost, making the effective price of the asset $\bar{p} \equiv \min\{B, p+\phi\}$. All assets mature according to a Poisson process with rate β and pay off R per unit when they mature. Investors also choose the scale of investments. They can invest with their cash only or borrow from the repo market against the investments, increasing the scale of investment.

When borrowing, investors take the terms of a repo contract (the haircut h, the repo spread s and the repo maturity rate σ) as given and solve for the optimal

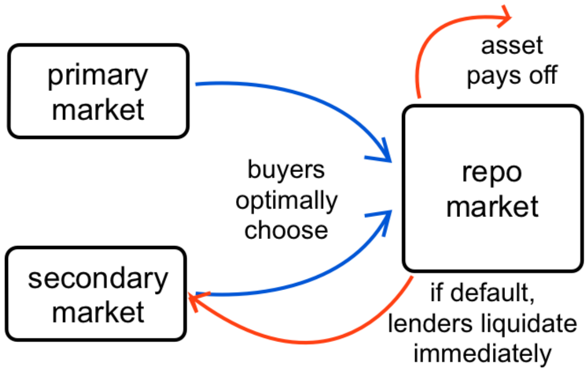


Figure I.2: Three Markets

leverage. They choose how much of their cash to pledge (haircut capital x) or to reserve (buffer capital $K-x$). When pledging x, the investor borrows $x(1/h-1)$ and scales the investments by x/h. The buffer capital $K-x$ is used to pay a stream of repo interest $(\rho+s)x(1/h-1)$. If investors run out of buffer capital before their investments mature, they are forced into bankruptcy and exit the system, and repo investors immediately liquidate the collateral in the secondary market. I assume repo investors protect themselves from bankruptcy by setting the haircut equal to the cost of liquidation. If investors are solvent when their investments mature, they receive R per unit from the investment, pay back the repo principal to repo investors, consume the rest and retire from the system. Whenever FIs exit, the system is replenished with the same mass of new FIs, fixing the total mass at 1.

Three markets are present in the model: (a) the primary market, where assets are initially traded; (b) the secondary market, where liquidated collateral from bankrupt FIs is traded; and (c) the repo market, where investors obtain financing from repo investors and collateral circulates (see Figure I.2). I assume the supply of the assets in the primary market is perfectly elastic, and focus on collateral circulation in the repo market and its interaction with the secondary market.

Without loss of generality, I set $K = 1$ and $B = 1$. I further assume the three Poisson processes for opportunity arrival, investment maturity and repo maturity are independent. The exogenous parameters of the economy are repo liquidity q, the rates of the three Poisson processes (α, β and σ), the riskless rate and time discount parameter ρ, the per-unit payoff from investment R and the transaction cost in the secondary market ϕ. A stationary equilibrium solves for the repo spread s^*, the optimal haircut capital x^*, the secondary market price p^*, the haircut h^*, and the masses of asset, repo, and cash investors in the financial system μ_I^*, μ_R^* and μ_C^* such that all FIs make optimal decisions and all markets clear.

Assets There is one class of assets that are risky, profitable and illiquid. The riskiness of assets lies in their random maturity. For an investment opportunity to be meaningful, the assets should have positive net present value. Thus, I assume that the net payoff is greater than the time value of money at the expected maturity ($1/\beta$):

$$R - 1 > \frac{\rho}{\beta}. \tag{I.1}$$

Repo contracts A repo (sale and repurchase agreement) is a form of collateralized lending with two main distinctions: (1) repos are often overcollateralized (the difference between the value of collateral and the principal of the loan as a fraction of the value of collateral is called a haircut); and (2) repo lenders, unlike other creditors, are exempt from automatic stay, meaning they can immediately liquidate collateral upon the borrower's bankruptcy. Section 4.3 explains institutional details relevant to the model's implications.

All borrowers and lenders meet and clear the repo market instantly. There are

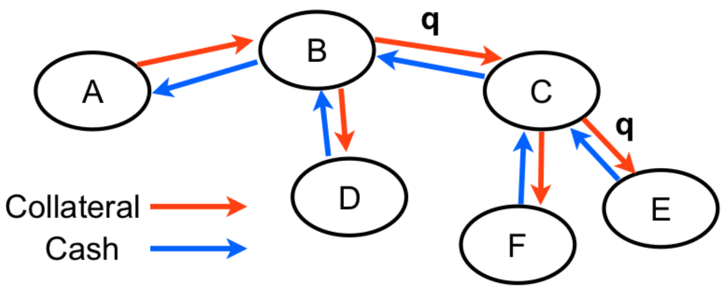


Figure I.3: Flow of Cash and Collateral

no search and matching frictions, and typically a pool of lenders are matched with a borrower. The equilibrium repo spread is determined by clearing the repo market. The spread is assumed to be the same across all FIs for tractability. That is, lenders cannot observe or verify the individual borrower's balance of buffer capital, although a low balance may shorten the maturity. The haircut is set so that repos are riskless from lender's perspective as in Geanakoplos (1997, 2010).[8] To simplify the wealth distribution, I further assume repo investors and cash investors continuously consume their interest income.

Collateral Circulation I use mismatch in the preferred timing of lenders and borrowers to generate collateral circulation, as in Kiyotaki and Moore (2001). Consider the borrower A and the lender B in Fig I.3. A pledges the investment as collateral and borrows cash from B. Until either A's investment matures or A runs out of cash and goes bankrupt, A keeps rolling over the repos at terms that are updated frequently. The lender B may be willing to stay as a lender to A until an opportunity arrives to the lender B. Then the lender B should rehypothecate (i.e., circulate) A's collateral to take the investment opportunity. B can circulate

[8] Whether this type of contract is optimal is beyond the scope of this paper. Rampini and Viswanathan (2010) show when there is limited enforcement, an optimal contract limits debt capacity to the amount that lenders can seize from collateral. I speculate that a similar setup might make the repo contract in this model optimal.

A's collateral to C with an exogenous probability q and use the cash to borrow from D against B's own investment. In a similar fashion, A's collateral may be circulated from C to E, and so on. Thus, repo liquidity q determines the extent to which a given collateral is circulated through the financial system.

Interpretation of the Setup A real-world example for the modeled assets is the originate-to-distribute model, where FIs lend to homeowners, pool the (subprime) mortgages, season the pool of mortgages, then securitize and sell them to buyers. The mortgages tend to be seasoned, typically for six months to one year, before they are sold as securities. This practice implies the pools of mortgages initially originated are illiquid, and this illiquidity is one of the key characteristics that the model captures in two ways. First, there is a transaction cost when the assets are traded in the secondary market before they mature. Second, only the financial institutions in the system can provide financing.

The key setup of the model is that FIs need an investment opportunity to purchase assets in either the primary or the secondary market. This investment opportunity is a modeling device that creates interim heterogeneity among institutions, and is similar in spirit to the slow moving capital of Duffie (2010); Mitchell, Pedersen, and Pulvino (2007). There are several other simplifying assumptions in the model including (1) the institutions cannot raise external equity either within the system or outisde the system; (2) the haircut is set so that repos are free of credit risk; (3) investment opportunities are specific to whom they arrive and cannot be transfered to other institutions; and (4) financial institutions cannot merge, or acquire other institutions. While relaxing these assumptions may be interesting topics for future research, the main implication of the model is robust that the cost of financing

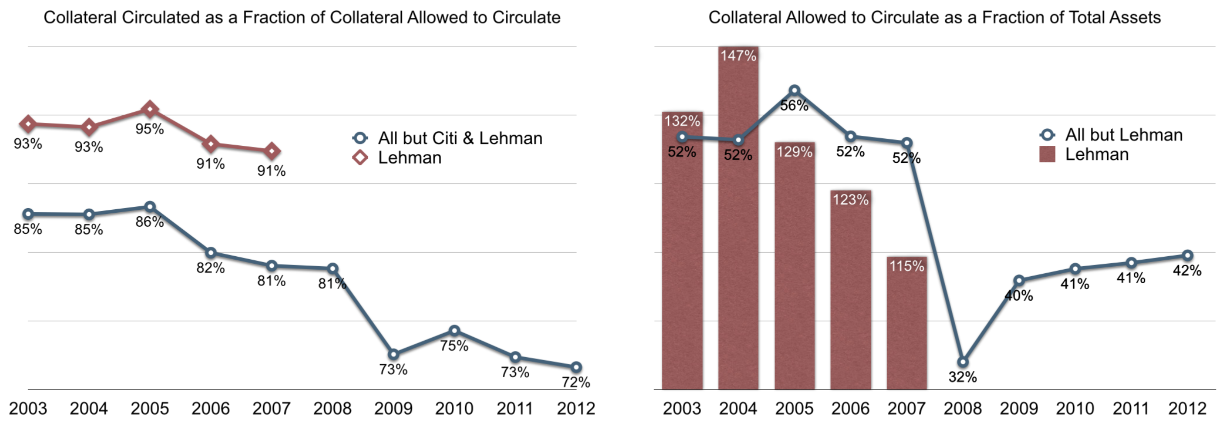


Figure I.4: The data is from 10-K reports of major US broker-dealers. The left figure excludes Citigroup as they report the value of collateral allowed to circulate only.

reflects the lost profit opportunities when repos are not perfectly liquid.[9]

The repo maturity is assumed random. The random maturity is common in a continuous time model such as Carr (1998) and He and Xiong (2012a,b), and makes the model tractable. Moreover, the random maturity captures the uncertainty in timing that lenders and borrowers get their cash and collateral back. The relationship between borrowers and lenders tends to be stable. In practice, even when repos have overnight maturity, they are not expected to be terminated in one day.[10] The observed short maturity should be understood as frequent updates in the terms of contracts (such as haircuts and repo spreads). In the model the repo maturity is random, but the terms are updated continuously.

While the model focuses on the repo markets, collateral circulation is not unique to the repo markets. The model can be applied to the broader financial markets. Repo liquidity, q, captures three aspects of real-world collateral circulation, and

[9] In Extensions I relax the assumption that FIs exit after their investments mature.

[10] In private conversation, practitioners mentioned that it is customary to give a warning in advance (one week to one month) when the parties want to terminate the contracts in bilateral repos. While most tri-party repos have very short overnight maturity, there is no aggregate data on the distribution of maturity in bilateral repos. Copeland, Martin, and Walker (2011) shows the actual relationship between borrowers and lenders tends to be stable in tri-party repos.

appears to have contracted during the recent crisis. First, borrowers who worry about getting their collateral back may not allow lenders to circulate it as the plot of collateral available for circulation (Figure I.1) shows. Second, lenders may not circulate collateral for precautionary motives. The proportion of collateral circulated relative to collateral allowed to circulate also declined in 2008 as in the left plot of Figure I.4. Third, the relationships among financial institutions depend on their network linkages; the ability of one institution to quickly circulate collateral is determined by other institutions that it is linked to. For example, the failure of Lehman Brothers greatly affected hedge funds whose sole prime broker was Lehman Brothers. The right plot of Figure I.4 shows that collateral circulation was much more important for Lehman Brothers than other major broker-dealers.[11]

3 Stationary Equilibrium

This section solves for a stationary equilibrium of a dynamic competitive economy. First I describe and solve the individual optimization problems of FIs. Given the individual solutions, I solve a stationary equilibrium by clearing the markets, and provide comparative statics.

3.1 Individual Optimization Problem

FIs with Investment Opportunities An FI with an investment opportunity ("asset investor") makes two decision: (1) he chooses to purchase assets in the primary or secondary market or both, and (2) he chooses whether and how much to borrow from the repo market. Since the primary market price is normalized to 1, the

[11] See also Aragon and Strahan (2012).

effective cost of asset per unit is

$$\bar{p} \equiv \min\{1, p+\phi\} \tag{I.2}$$

The investor chooses whether or not to borrow against the assets to increase the scale of his investment. The value of unlevered investment is

$$V_{unlevered} \equiv E\left\{e^{-\rho\tilde{\tau}_\beta}\frac{R}{\bar{p}}\right\} = \left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}}, \tag{I.3}$$

where $\tilde{\tau}_\beta$ is the random maturity of the asset. The unlevered value is greater than 1 by the profitability assumption (I.1).

If the investor chooses to borrow from the repo market, he needs to decide how much to borrow. The investor can increase scale of the investments if he pledges more. On the other hand, the investor has to pay the repo interest to stay solvent. Thus, he needs to set aside cash whose balance he keeps in his checking account, earning the riskless rate of ρ.[12] Suppose the investor pledges x units of haircut capital and sets aside $1-x$ units of buffer capital. Then, the investor can borrow $\left(\frac{1}{h}-1\right)x$ to invest

$$\frac{x}{h} = x + \left(\frac{1}{h}-1\right)x. \tag{I.4}$$

Per unit of haircut capital pledged, the investor can buy $\frac{1}{h\bar{p}}$ units of the assets from either the primary or the secondary market, whichever he prefers. When the investment matures, each unit of the asset pays off R and he pays back the

[12]Note that the investor keeps his buffer capital in the checking account and earns ρ. The investor is excluded from lending the buffer capital in the repo market because he needs to pay the interest continuously.

principal. Thus the leveraged investment per unit haircut capital pays off

$$\bar{R} \equiv \frac{R}{h\bar{p}} - \left(\frac{1}{h} - 1\right) = 1 + \frac{1}{h}\left(\frac{R}{\bar{p}} - 1\right), \tag{I.5}$$

whose net payoff is simply the asset's net payoff multiplied by leverage $1/h$.

The investor pays the flow repo interest out of his buffer capital $1 - x$. The repo interest is the repo rate $(\rho + s)$ multiplied by the principal. Thus, the investor remains solvent until he exhausts his buffer capital at $T_\delta(x)$, which solves

$$\int_0^{T_\delta(x)} e^{-\rho\tau} (\rho + s) \left(\frac{1}{h} - 1\right) x d\tau = 1 - x. \tag{I.6}$$

The duration of investor's solvency, $T_\delta(x)$, decreases in haircut capital x because of the reduced buffer capital as well as the increased repo interest. Then the default probability of an investor can be obtained as

$$\theta(x) \equiv Pr\left\{\tilde{\tau}_\beta > T_\delta(x)\right\} = e^{-\beta T_\delta(x)}. \tag{I.7}$$

Combining (I.5), (I.6), and (I.7) yields the value of the leveraged investment. For emphasis, I write it as a function of the repo spread s as below:

$$V_{leveraged}(s \mid h, p, \beta, R) \equiv$$

$$\max_x \left[x(1 - \theta(x)) E\left\{ e^{-\rho\tilde{\tau}_\beta}\bar{R} + \int_{\tilde{\tau}_\beta}^{T_\delta(x)} e^{-\rho\tau} \left(\frac{1}{h} - 1\right)(\rho + s) d\tau \mid \tilde{\tau}_\beta \le T_\delta(x) \right\} \right], \tag{I.8}$$

where the first part $(x(1 - \theta x))$ is the haircut capital, which determines the scale of investment, multiplied by the probability that the investor will remain solvent; and the second part is the expected payoff, conditional on solvency, per unit haircut capital, which is the sum of the investment payoff and the remaining buffer capital. The first order condition with respect to haircut capital x when the leveraged investment is profitable is given by[13]

$$\underbrace{\left\{\frac{\beta\bar{R}-\left(\frac{1}{h}-1\right)(\rho+s)}{\rho+\beta}\right\}(1-\theta(x^*))-1}_{\text{marginal benefit}}=\underbrace{-\frac{dT_\delta}{dx^*}\left\{\beta\bar{R}-\left(\frac{1}{h}-1\right)(\rho+s)\right\}x^*}_{\text{marginal cost}}, \quad \text{(I.9)}$$

where the L.H.S. is the marginal benefit of haircut capital and the R.H.S. is the marginal cost of haircut capital. While the extra haircut capital increases the scale of investment, it shortens the time that an investor would remain solvent, making them more likely to default. Thus, the optimal level of haircut capital x^* (thus the optimal leverage x^*/h) trades off the increased scale with the possibility of bankruptcy.

Using the first order condition (I.9), we can write the leveraged value function as

$$V_{leveraged}(s)\equiv 1+\left[\frac{\beta\bar{R}}{\left(\frac{1}{h}-1\right)(\rho+s)}-1\right]\theta^*(s), \quad \text{(I.10)}$$

where the optimal default probability $\theta^*(s)\in(0,1)$ is the unique solution to

$$\left(\frac{1}{\rho}+\frac{1}{\left(\frac{1}{h}-1\right)(\rho+s)}\right)\theta^*-\frac{\beta}{\rho(\rho+\beta)}\theta^{*\frac{\rho+\beta}{\beta}}=\frac{1}{\rho+\beta}-\frac{1}{\beta\bar{R}-\left(\frac{1}{h}-1\right)(\rho+s)}. \quad \text{(I.11)}$$

Finally, the investor compares the value of leveraged and unlevered investment and chooses whether or not to borrow. While the value of unlevered investment is independent of the repo spread, the leveraged investment decreases in the repo

[13] If $V_{leveraged}\leq 1$, the optimal haircut capital is simply $x^*=0$ and $V_{leveraged}=1$.

spread. There is a unique upper threshold of the repo spread at which the investor would stop borrowing from the repo market. The lemma below summarizes the result. All proofs are in the Appendix.

Lemma 1. *The value function of an asset investor is given by*

$$V_I(s) = \begin{cases} V_{leveraged}(s) = 1 + \left[\frac{\beta \bar{R}}{\left(\frac{1}{h}-1\right)(\rho+s)} - 1\right]\theta^*(s) & \text{if } s < \hat{s} \\ V_{unlevered} = \left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}} & \text{otherwise} \end{cases}, \tag{I.12}$$

where $\theta^*(s)$ solves (I.11). That is, an investor prefers to borrow from the repo market if the repo spread is lower than $\hat{s}$ and not to borrow otherwise. The cutoff $\hat{s}$ is always positive and uniquely determined by

$$\theta^*(\hat{s}) = \left[\frac{\beta \bar{R}}{\left(\frac{1}{h}-1\right)(\rho+\hat{s})} - 1\right]^{-1} \left[\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}} - 1\right]. \tag{I.13}$$

FIs without Investment Opportunities FIs without an opportunity choose whether to invest in cash or to invest in repos. A cash investor continuously earns and consumes the riskless interest of rate ρ. Once an investment opportunity arrives at τ_α, he instantly becomes an asset investor. Thus, the value of a cash investor is

$$V_C = E\left\{\int_0^{\tilde{\tau}_\alpha} e^{-\rho\tau}\rho d\tau + e^{-\rho\tilde{\tau}_\alpha} V_I\right\} = 1 + \frac{\alpha}{\rho+\alpha}(V_I - 1). \tag{I.14}$$

A repo investor continuously earns and consumes the repo interest at rate $\rho + s$ until one of three events occurs: (1) an investment opportunity arrives to the repo investor at τ_α, (2) the borrower's investment matures at τ_β or (3) the repo matures at τ_σ.[14] If an opportunity arrives first, the repo investor can circulate

[14]The independence assumption of the processes guarantees no two events occur simultane-

collateral with probability q, get his cash back, and become an asset investor. If an opportunity arrives later than the other events, the repo investor chooses whether to continue to lend or to become a cash investor.[15] Then the value of a repo investor can be written as

$$V_R = E\left\{\int_0^{\tilde{\tau}_m} e^{-\rho\tau}(\rho+s)\,d\tau + e^{-\rho\tilde{\tau}_m}\left[V_R + q(V_I - V_R)_{\{\tilde{\tau}_m=\tilde{\tau}_\alpha\}} + \max\{V_C - V_R, 0\}_{\{\tilde{\tau}_m\neq\tilde{\tau}_\alpha\}}\right]\right\}, \tag{I.15}$$

where I denote by $\tilde{\tau}_m \equiv \min\left\{\tilde{\tau}_\alpha, \tilde{\tau}_\beta, \tilde{\tau}_\sigma\right\}$.

Since FIs can always choose to invest in cash, it holds that $V_R \geq V_C$ in equilibrium. Thus, the value of a repo investor can be simplified as

$$V_R = 1 + \frac{s}{\rho+q\alpha} + \frac{q\alpha}{\rho+q\alpha}(V_I - 1), \tag{I.16}$$

which is independent of the repo maturity rate because of the independence assumption.

While the value of a cash investor is independent of the repo spread, the value of a repo investor increases in the repo spread. Thus, there is a unique lower threshold of the repo spread $s_l \geq 0$ at which FIs without opportunities would start lending in the repo market.

Lemma 2. *For a given level of the investor's value (V_I), there is a minimum repo spread*

$$s_l(V_I) = (1-q)\left(\frac{\alpha\rho}{\rho+\alpha}\right)(V_I - 1),$$

ously.

[15] If the borrower goes bankrupt, repo investors immediately liquidate the collateral in the secondary market. Since the haircut is set so that repos are riskless, the repo investors can recover the full principal.

at which FIs without investment opportunities choose to invest in repos.

The minimum repo spread is nonnegative, and strictly positive if and only if repos are not perfectly liquid ($q < 1$). The spread compensates repo investors for the potential opportunity cost of lending. The opportunity cost arises from not being able to circulate collateral with probability $1-q$ and having to forego profitable investments ($V_I - 1$).

Recall that the value of an investor V_I depends on the repo spread (I.12). With a slight abuse of notation, I denote by s_l the solution to the fixed point problem[16]

$$s_l = (1-q)\left(\frac{\alpha\rho}{\rho+\alpha}\right)(V_I(s_l)-1). \tag{I.17}$$

Assets versus Repos So far I have assumed FIs take an investment opporuntity when they can. Asset investors, however, are free to become repo or cash investors. Intuitively, investors would never prefer to become cash investors, but they may prefer to become repo investors if the repo spread is sufficiently high. From (I.16), we find that the investment opportunity is more attractive than investing in repos (i.e., $V_I > V_R$) if and only if the repo spread is s is less than

$$\hat{\hat{s}} = \rho\left(V_I\left(\hat{\hat{s}}\right)-1\right). \tag{I.18}$$

Note that $\hat{\hat{s}}$ is always greater than the minimum spread for lending s_l, as the investor value is strictly higher than the value of a cash investor, regardless of repo liquidity or the arrival rate of opportunities.

Combining (I.12) and (I.18), we find that an investor chooses to borrow if and only

[16]The solution exists and is unique as $V_I(\cdot)$ is decreasing in s.

if the repo spread is lower than both thresholds $\hat{s}$ and $\hat{\hat{s}}$. Therefore, the maximum repo spared for borrowing is given by

$$s_b \equiv \min\left\{\hat{s}, \hat{\hat{s}}\right\}. \tag{I.19}$$

3.2 Equilibrium Solution

Equilibrium Repo Spread FIs without investment opportunities will only invest in repos if the repo spread is higher than s_l (I.17), and asset investors will only borrow if the repo spread is lower than s_b (I.19). Thus, for the repo market to open, the equilibrium repo spread s^* satisfies

$$s^* \in [s_l, s_b]. \tag{I.20}$$

At any repo spread $s \in [s_l, s_b]$, repo contracts create a surplus. How repo investors and asset investors divide the surplus depends on their bargaining power, which in turn is determined by the market clearing condition. The bargaining problem between borrowers and lenders is described in detail in the Appendix. Here I focus on the case where all the surplus of repos is extracted by borrowers (asset investors) and the equilibrium repo spread equals the lender's minimum s_l. Proposition 6 provides the condition for which this is the case.[17]

If the minimum spread for lending s_l is higher than the maximum spread for borrowing s_b, repos do not create surplus and the repo markets shut down. Then the shadow price for the repo spread equals s_l at which no investors borrow.

[17] I find the inherent illiquidity of asset, transaction cost parameter ϕ, plays a key role.

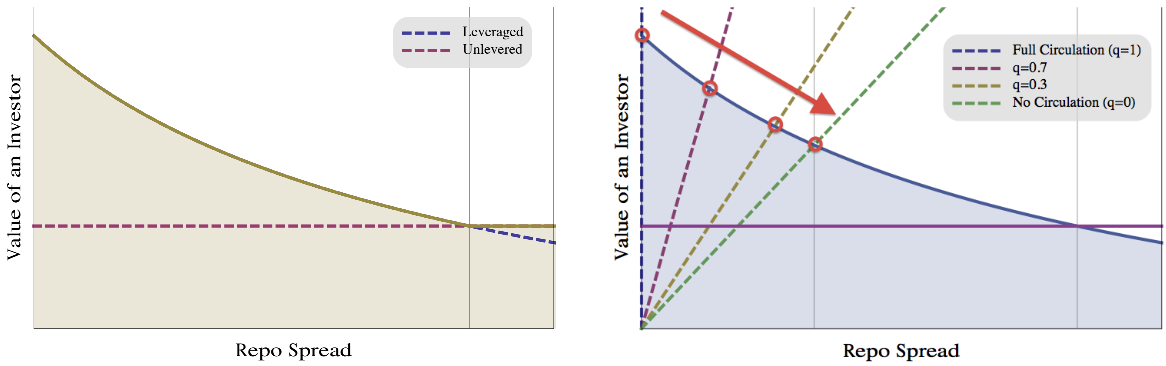


Figure I.5: The Fixed-Point Problem

Therefore, the equilibrium repo spread s^* is given as

$$s^* = (1-q)\left(\frac{\alpha\rho}{\rho+\alpha}\right)(V_I^* - 1), \tag{I.21}$$

where $V_I^* = V_I(s^*)$. Figure I.5 illustrates how the equilibrium spread s^* and the value of an investor V_I^* are determined simultaneously as solutions to the fixed-point problem. While the value of an investor decreases in the repo spread, the indifference repo spread increases in the value of an investor. Equilibrium values are found where the two meet for given level of repo liquidity.

Equilibrium Composition of the Financial System Now I solve for equilibrium masses of different types of FIs: investors (μ_I), repo investors (μ_R) and cash investors (μ_C). While an individual FI can transition between different types, the composition of the financial system remains constant. The masses, μ_I, μ_R and μ_C, are solutions to the three equilibrium conditions. First, the total mass of FIs is fixed at unity, as all the exiting FIs are replaced by new FIs.

$$\mu_I + \mu_R + \mu_C = 1 \tag{I.22}$$

Second, the inflow and outflow of investors should be equal for stationarity. The inflow of investors is determined by the arrival of opportunities. A new investor might have been a cash investor or a repo investor who circulated collateral. The outflow of investors is determined by the asset maturity and bankruptcy.

$$\underbrace{\alpha\left(\mu_C + q\mu_R\right)}_{\text{inflow}} = \underbrace{\theta^*\left(s^*\right)\alpha\left(\mu_C + q\mu_R\right) + \beta\mu_I}_{\text{outflow}} \tag{I.23}$$

Lastly, the mass of repo investors is determined by clearing the repo market.

$$\mu_R = \left(\frac{1}{h} - 1\right) x^*\left(s^*\right)\mu_I \tag{I.24}$$

Using (I.22), (I.23), and (I.24), we find equilibrium masses, μ_I^*, μ_R^*, and μ_C^*, as in (47) - (51).

Equilibrium Haircut Since investors can choose to purchase assets from the primary market at a price of 1 and the secondary market at a price of p, the equilibrium secondary market price p^* cannot be higher than $1-\phi$. On the other hand, the supply of assets in the secondary market at any point in time is strictly lower than the total demand for the assets. This is because the optimal leverage is constant across investors, and the inflow of investors is strictly greater than the outflow of investors due to bankruptcy (I.23). As some incoming investors must purchase assets from the primary market, p^* cannot be lower than $1-\phi$. Therefore, we have

$$p^* = 1 - \phi. \tag{I.25}$$

The equilibrium haircut h^* is simply set so that repo investors are protected from the cost of immediate liquidation. From the equilibrium secondary market price above, we have

$$h^* = \phi. \tag{I.26}$$

So we have found a competitive stationary equilibrium. It always exists and is unique.

Proposition 1. *There exists a unique competitive stationary equilibrium for a given set of parameters* $\{q, \alpha, \beta, \sigma, \rho, R, \phi\}$. *The equilibrium is a tuple* $\left(s^*, x^*, h^*, p^*, \{\mu_i^*\}_{i\in\{I,R,C\}}\right)$ *that satisfies (61), (47) - (51) , (I.25), (I.26),*

$$x^* = \theta^{-1}(\theta^*) \tag{I.27}$$

where $\theta(\cdot)$ *and* θ^* *are given in (I.7) and (I.11), if* $s^* \leq s_b$ *and* $x^* = 0$*, otherwise.*

Note that the equilibrium repo spread s^* is positive whenever collateral circulation is frictional ($q < 1$). Specifically, the equilibrium spread is no less than the minimum spread of lending s_l, which is positive if and only if collateral chains are frictional. This is the case even though haircuts are set so that repos are free of credit risk, meaning repo investors can recover the full principal regardless of borrower's bankruptcy. Thus the repo spread in this model is not compensation for credit risk but compensation for the opportunity cost that repo investors incur. With frictional collateral chains, repo investors face a positive probability ($1 - q > 0$) that they may get an investment opportunity, but cannot circulate collateral and have to forego profitable investments. The opportunity cost, and thus the equilibrium spread, increase in the rate of opportunity arrival α, frictions

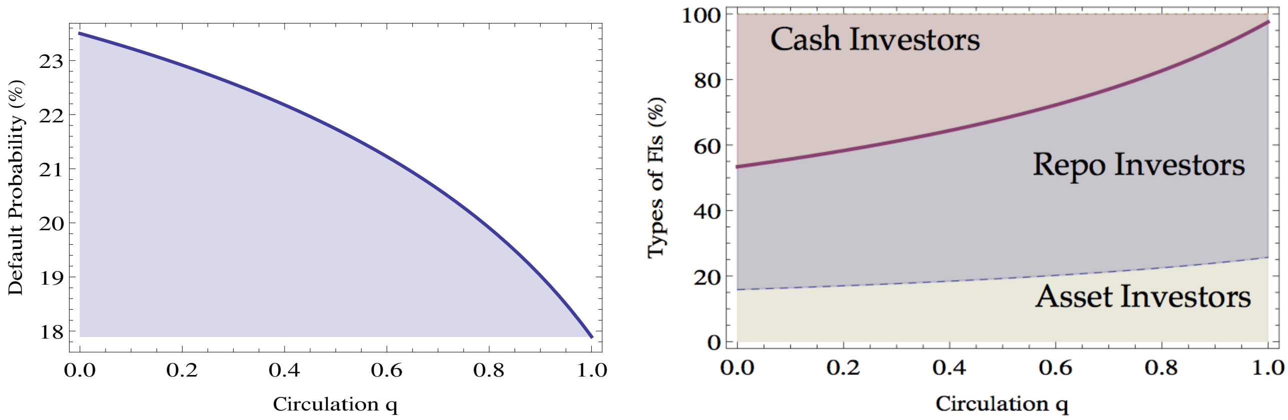
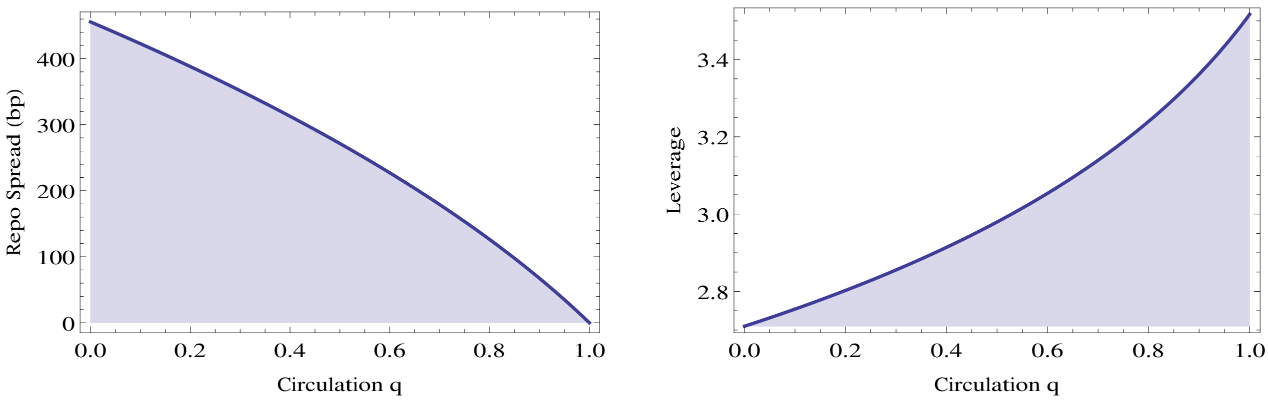


Figure I.6: Comparative Statics. The baseline parameters are provided in Section 4.

in collateral chains $(1-q)$, and the profitability of investments $(V_I - 1)$. Since illiquid investments must be financed within a system that has limited cash, the cost of capital may be higher than that for liquid investments, captured by the repo spread.



Comparative Statics Under conditions characterized in Proposition 2, I find that collateral circulation increases aggregate investment, exhibiting *three* multiplier effects. As repos become more liquid the repo spread decreases and thus (1) individual FIs optimally take higher leverage and undertake a larger scale of investment; (2) their assets are less likely to default, i.e., are safer, even with higher leverage;

and (3) there are more FIs that are investing in their own assets in the system.[18]

Proposition 2. *Provided that the transaction cost ϕ satisfies (62), and the asset payoff R satisfies*

$$R \geq 1 + 4\frac{\rho}{\beta}, \tag{I.28}$$

we have $\forall q \in [0,1]$,

$$\frac{ds^*}{dq} \leq 0, \ \frac{d\theta^*}{dq} \leq 0, \ \frac{dx^*}{dq} \geq 0 \ \textit{and} \ \frac{d\mu_I^*}{dq} \geq 0. \tag{I.29}$$

Thus, we answer the first question posed in the introduction. Clearly, collateral circulation not only creates double-counting problems, but also creates real multiplier effects. The analogy between circulating repo collateral and the money multiplier, such as in Gorton and Metrick (2011) and Shin (2010), appears to be accurate. Multiplier effects of collateral circulation operate through the endogenous cost of financing, which in turn affects optimal leverage, default probability, and the mass of investors.

4 Shocks and Fragility

Now I introduce various shocks to the economy and study transition paths. First I explain the equilibrium concepts. Second, I compute the transition paths when the economy is subject to various shocks. Third, I explain the positive feedback effect that arises from endogenous haircuts and repo spreads. Lastly, I connect the model's implications to the financial crisis of 2008.

[18]In general, the default probability θ^* is concave in the repo spread, at first increasing and then decreasing. However, I find that as long as the payoff of unit investment R satisfies (I.28), the equilibrium repo spread always lies in the region such that the default probability is increasing in the repo spread, regardless of other parameters.

4.1 Deterministic Transition Path

Consider an economy with a set of parameters $\{q_0, \alpha_0, \beta_0, \sigma_0, \rho_0, R_0, \phi_0\}$. By Proposition 1, there exists a unique stationary equilibrium denoted by $(x_0^*, s_0^*, h_0^*, p_0^*, \mu_0^*)$. Suppose a shock that is completely unanticipated and permanent arrives at $t = T$. The shock is defined as changes in the parameters to $\{q, \alpha, \beta, \sigma, \rho, R, \phi\}$, where q is the collateral circulation parameter, α, β, σ are the intensities of the investment opportunity arrival, investment maturity and repo maturity, ρ is the time discount parameter and the riskless rate, R is the payoff of the investment at maturity, and ϕ is the transaction cost of secondary market assets.

I study a competitive equilibrium along a deterministic transition path. Once a shock arrives, all FIs in the economy have full information about the new parameters instantly. Moreover, all FIs have perfect foresight with respect to the dynamics of equilibrium prices and terms of repo contracts. All FIs make individually optimal decisions and markets clear at all instants. There is no remaining uncertainty, making the transition path deterministic. As discussed in Section 2, I let the terms of repos - both the repo spread and haircut - adjust continuously regardless of the random repo maturity, to reflect the reality of short repo maturities.

Definition 1. A competitive equilibrium along a deterministic transition path with a shock at T is a tuple $\{X_t, S_t, H_t, P_t, \mu(t) \mid \forall t \in [T, \infty)\}$ that satisfies

(i) $\forall t \geq T$, FIs with an opportunity choose the optimal level of leverage (including completely refraining from borrowing) and pledge X_t if borrowing, taking as given the future repo rate $\{S_\tau\}$, the haircut $\{H_\tau\}$, and the price of secondary market assets $\{P_\tau\}$ $\forall \tau \geq t$. FIs without opportunities optimally choose either to lend in the repo market or to wait.

(ii) $\forall t \geq T$, all markets clear or the price hits the boundary ($P_t = 0$).

(iii) The mass of investors, repo investors and cash investors ($\mu^*(t) = (\mu_I^*(t), \mu_R^*(t), \mu_C^*(t))$, $\forall t \geq T$) evolve according to processes consistent with FIs' optimal strategies.

First consider an intuitive case. One candidate for a transition path is a stable path where the secondary market price remains constant ($P_t \equiv 1 - \phi$). Recall that in a stationary equilibrium the demand for assets is strictly higher than the supply of the secondary market. That is, the secondary market has a capacity to absorb some extra supply of liquidated assets without affecting its price. As the secondary market price remains constant, so does the haircut. Fixing the haircut, the optimal choices of FIs after the shock are similar to those in a stationary equilibrium with new parameters.

Definition 2. A stable path is a competitive equilibrium along a deterministic transition path such that for $\forall t \geq T$,

$$X_t = x_1^*,\ S_t = s_1^*,\ H_t = \phi \text{ and } P_t = 1 - \phi \tag{I.30}$$

where $(x_1^*, s_1^*, h_1^* = \phi, p_1^* = 1 - \phi)$ are the values of a stationary equilibrium with the new parameters $\{q, \alpha, \beta, \sigma, \rho, R, \phi\}$.

Here I describe the dynamics of the masses of different types of FIs for the case when both the stationary equilibrium before the shock and after the shock are of Type 1 (i.e., all investors borrow). Even though all the prices instantly jump and stay constant after the shock, the masses of different types of FIs ($\mu^*(t) =$

$(\mu_I^*(t), \mu_R^*(t), \mu_C^*(t))$ for $\forall t \geq T$) slowly converge to the new stationary equilibrium levels. FIs enter and exit the system and transition between types at different rates from those before the shock. Moreover, immediately after the shock, FIs who entered before and after the shock coexist. Once all the old FIs exit the system, rates at which FIs enter and exit the system and transition between types are the same as those in a stationary equilibrium with new parameters, hence converging to the stationary equilibrium levels and remaining constant afterwards.

Consider an FI whose project started at t_0, before the shock, and is ongoing at the time that the shock arrives (i.e., $t_0 < T$ and $t_0 + T_{\delta 0}^* > T$, where $T_{\delta 0}^* \equiv T_\delta(x_0^*)$ and $T_\delta(\cdot)$ is defined in (I.6)). If it were not for the shock, the FI would either succeed in their investment and retire from the system or go bankrupt by $t_0 + T_{\delta 0}^*$. However, with the shock the repo spread changes from s_0^* to s_1^* although the haircuts remain the same along the stable path. Therefore, the duration that the FI can remain solvent changes to $\tilde{T}_\delta(t_0)$ where $\tilde{T}_\delta(\cdot)$ solves[19]

$$(\rho_0 + s_0^*) \int_T^{t_0 + T_{\delta 0}^*} e^{-\rho\tau} d\tau = (\rho + s_1^*) \int_T^{t_0 + \tilde{T}_\delta(t_0)} e^{-\rho\tau} d\tau \tag{I.31}$$

That is, if the shock increases the repo spread ($s_1^*/s_0^* > 1$), the FI can remain solvent for a shorter horizon ($\tilde{T}_\delta(t_0) < T_{\delta 0}^*$). It is more sensitive to the shock if their investment started shortly before the shock.

By substituting $t_0 \to T$ into (I.31), we can see that FIs who started investments right before the shock can remain solvent until $\hat{T}$, after which all the old FIs exit

[19](I.31) can be simplified as $T_\delta^*(t_0) = T - t_0 - \frac{1}{\rho} \log \left[\frac{s_1^* - s_0^*}{s_1^*} + \frac{s_0^*}{s_1^*} e^{-\rho\left(t_0 + T_{\delta 0}^* - T\right)} \right]$.

the system.

$$\hat{T} \equiv T + \tilde{T}_{\delta}(T_0) = T - \frac{1}{\rho}\log\left[\frac{s_1^* - s_0^*}{\rho + s_1^*} + \left(\frac{\rho_0 + s_0^*}{\rho + s_1^*}\right)e^{-\rho T_{\delta 0}^*}\right] \tag{I.32}$$

which reduces to $T + T_{\delta 0}^*$ if it were not for the shock.

Between the arrival of the shock at T and the exit of all old FIs at $\hat{T}$, the mass of investors $\mu_I^*(t)$ includes both new and old FIs and evolves as below. $\forall t \in \left[T, \hat{T}\right]$,

$$\frac{d\mu_I^*(t)}{dt} = \underbrace{\alpha\left(\mu_W^*(t) + q\mu_L^*(t)\right)}_{\text{inflow}} - \left[\underbrace{\beta\mu_I^*(t) + \Theta^*(t)\cdot\alpha_0\left(\mu_{0W}^* + q_0\mu_{0L}^*\right)}_{\text{outflow}}\right] \tag{I.33}$$

The inflow to the investor type is determined by the new circulation parameter q and the masses of repo investors and cash investors. The outflow from the investor type is determined by the new intensity of investment maturity β and the defaults of old investments. Here,

$$\Theta^*(t) \equiv \exp\left\{-\left[(\beta - \beta_0)(t - T) + \beta_0 T_{\delta}^*(t_0)\right]\right\} \tag{I.34}$$

represents the remaining investments that started at $t_0 (< T)$ and default at t (i.e., $t_0 + \tilde{T}_{\delta}(t_0) = t$, where $\tilde{T}_{\delta}(\cdot)$ is defined in (I.31)). Investments mature at rate β_0 before the shock and β after the shock. Without a shock to the investment maturity, $\Theta^*(t)$ is simply $\theta_0^* \equiv \theta(x_0^*)$ where $\theta(\cdot)$ is defined in (I.7). The dynamics of μ_L^* and μ_W^* follow.

The dynamics of $\mu_I^*(\cdot)$ after $\hat{T}$ are omitted as it is similar to that of a stationary equilibrium in (I.22) - (I.24) in the sense that it is solely determined by investors that entered after the shock, and their optimal choices are the same as those in the new stationary equilibrium. Initially, $\mu_I^*(\cdot)$ is not constant as it depends on the stocks of repo investors and cash investors that are different from that of the stationary equilibrium. Then all the masses converge to the stationary equilibrium (see I.7).

The outflow of investors due to bankruptcy in (I.34) determines the supply of the secondary market. As the haircut remains constant along a stable path, the scale of investment is proportional to the haircut capital x_0^* and x_1^*. Recall that for a stable path to exist, the secondary market must absorb potential increases in supply so that the prices do not fluctuate. Thus by comparing the demand and supply for assets we get the following necessary condition for the existence of a stable path.

Proposition 3. *A necessary condition for the existence of a stable path is for* $\forall t \in \left[T, \hat{T}\right]$,

$$\frac{\alpha_1 x_1^*}{\alpha_0 x_0^*}\left(\frac{\mu_C(t) + q\mu_R(t)}{\mu_C^* + q_0\mu_R^*}\right) > \Theta^*(t) \tag{I.35}$$

Thus, whether a given shock triggers a repo run. When repos have little moneyness the financial system hoards large amounts of cash. The cash reserves, although "inefficient" when compared to an equilibrium with more repo moneyness, make the financial system more resilient. When there is a shock, the cash reserves help buffer the shock along the transition path to the new equilibrium. When reserves are insufficient, a shock results in fire sales and a repo run.

Consider the case when the necessary condition (I.35) does not hold. That is, there is an excess supply in the secondary market due to the shock and the price has to adjust. Intuitively, the secondary market price should adjust downwards to attract investors to buy more assets. The price p of the asset in the secondary market is determined by market clearing conditions, similar to the cash-in-the-market pricing of Allen and Gale (1994, 2005).

One extreme example for a transition path is a repo run, where the secondary market price hits its lower bound ($P_T = 0$). In a repo run, collateral is worthless and all existing investors are forced into bankruptcy. All FIs become cash investors and the economy starts anew and slowly converges to a new stationary equilibrium. The following proposition shows that a repo run is indeed a competitive equilibrium transition path whenever (I.35) doesn't hold.

Proposition 4. *Define a repo run as a competitive equilibrium along a deterministic transition path such that*

$$H_T = 1,\ P_T = 0 \text{ and } \mu_W(T) = 1$$

If (I.35) is not satisfied, then a repo run exists. Moreover, the repo spread at the shock is

$$S_T = \frac{\alpha\rho}{\alpha+\rho}(1-q)\left[\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\phi} - 1\right] \tag{I.36}$$

4.2 Numerical Examples

Parameter selection I use parameter values which are consistent with the data and previous research. Following the 4.97% average one-year Treasury rate during the first half of 2007 (which is before the financial crisis), I set the discount rate to $\rho = 0.05$. The transaction cost ϕ equals the haircut in a stationary equilibrium. According to Copeland, Martin, and Walker (2011), the haircuts for private collateral in the bilateral repo market during a stable period were about 5% for high-grade corporate debt, 17% for Alt-A, prime MBS and 25% for subprime MBS, so I set $\phi = 0.15$. For the investment maturity intensity (β), I set $\beta = 0.5$, meaning it takes about two years to finish a project. I set $\alpha = 0.2$ and $R = 1.8$. The model is robust to a wide range of parameters. As a benchmark, I use frictionless collateral circulation ($q_0 = 1$). Based on I.1, I use $q = 0.5$ as the level of collateral circulation after the shock.

Iteration method Solving for a competitive equilibrium along a deterministic transition path is similar to solving a fixed point problem. Once the shock arrives, all FIs learn the parameters, future prices, and repo terms perfectly, and make an optimal decisions. In turn, all markets clear at those prices. I use an iteration method to find a transition path. That is, I start with various initial levels of prices and terms, solve for the optimal policies, and check whether the markets clear at those prices. The terms adjust accordingly when the markets do not clear. I keep iterating this procedure until the full paths solve the fixed point problem. The investors' optimization problems and the clearing of the secondary market when both the haircut and repo spread can fluctuate are described in Extensions.

Circulation shock First I study a negative shock to collateral circulation where the circulation parameter q decreases from the initial frictionless value $q_0 = 1$. Using $q = 0.5$, which was documented during the financial crisis, I find the economy is stable, with no changes in the secondary market price and thus the haircut. The repo spread instantly jumps to the new stationary equilibrium level (from 0 bp to 271 bp), while the masses of investors, repo investors and cash investors slowly converge to the new stationary equilibrium levels (Top left of I.7).

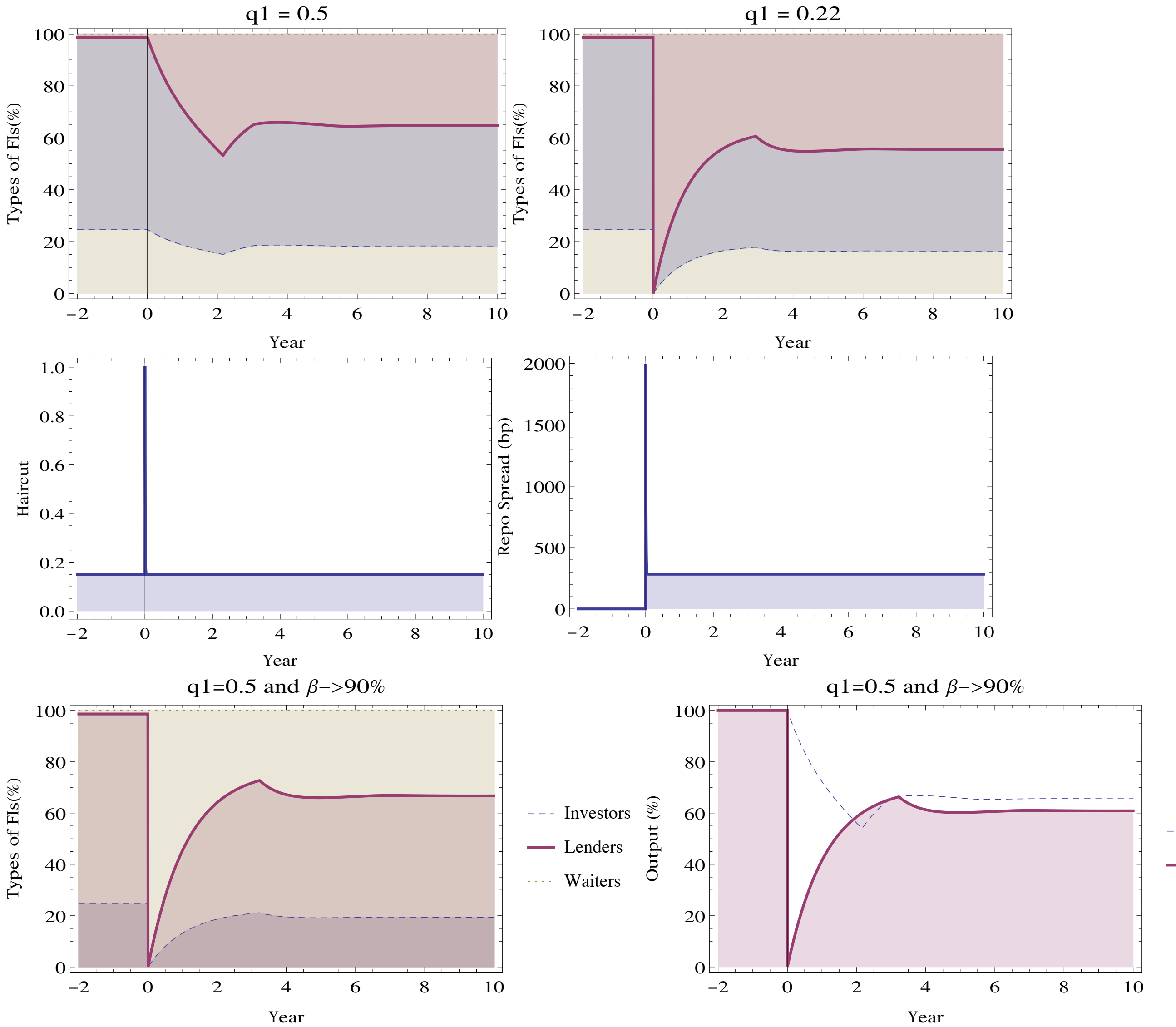


Figure I.7: Top figures report the evolution of masses of different types of FIs when the circulation parameter decreases from 1 to 0.5 (left) and 0.22 (right). Middle figures report the equilibrium haircut (left) and repo spread (right) when the economy exhibits a repo run. Bottom figures report the evolution of masses of different types of FIs (left) and the output as a fraction of the pre-shock level (right) when the economy suffers both the circulation shock ($q1$ =0.5) and the economic slowdowns (β decreases by 10%).

Numerically, I experiment by lowering q in 1% increments. From Proposition 3, I find that the stable path is no longer supported when $q \leq 0.22$. Once the circulation parameter hits the threshold 0.22, the secondary market cannot absorb the increase in supply and a fire-sale discount is necessary to clear the market. Then the powerful positive feedback effect, as explained more in detail below, starts to kick in. As a result, the iteration method with the initial condition of the stable path ends up converging to the run equilibrium. In I.7, I document the run equilibrium when $q = 0.22$. At the shock, the haircut increases to unity and all the existing borrowers are forced into bankruptcy. With the secondary market price at 0, the spread jumps up to 35.92% (which is the shadow spread at which no repo contracts occur), then immediately comes back to 381bp. While the individual optimization problem after the shock is similar to that of a stationary equilibrium, the measures of types of bankers slowly converge to the stationary equilibrium level (Top right of I.7). As shown, the measure of investors increases continuously from 0 at the shock, but then decreases for a short period of time before it converges to the stationary equilibrium level. The overshooting is coming from the lack of a stock of investment in the initial period. Once investments started after the shock start to go bankrupt , the measure decreases and converges to the stationary equilibrium level.

Other shocks I consider other types of shocks to the economy. First, I experiment with the investment maturity intensity declining by 10% ($\beta_0 = 0.5 \rightarrow \beta_1 = 0.45$) simultaneous with the circulation shock ($q_0 = 1 \rightarrow q_1 = 0.5$). The reduction in the maturity intensity implies it takes 10% longer for a project to mature, creating "economic slowdowns". I find that when the combined circulation shock ($q_1 = 0.5$)

and economic slowdown hits the economy, the economy does not support the stable path any longer and exhibits a repo run (middle and bottom figures of I.7). The haircut instantly jumps up to 1 and comes back to the stationary equilibrium level (0.15) immediately. The repo spread increases up to 25.67% then comes back to the new stationary equilibrium level 2.83%.

The economic slowdown further directly affects the output of the economy. The flow output of the economy is simply given by $\beta R(1-\mu_W(t))$, where all the capital that is utilized in the economy $(1-\mu_W(t))$ produces output R per unit at maturity, and matures at rate β. Thus, with the economic slowdown the economy produces lower output and takes longer to recover to the new stationary equilibrium. The bottom right figure of I.7 shows that the temporary impact of the shock can be much greater than the permanent impact when the shock causes a repo run.

Moreover, the reduction in the opportunity arrival intensity α, similar to the shock to β, exacerbates the instability of the economy. Recall that the stationary equilibrium repo spread at which repo investors and cash investors are indifferent increases in α (I.17). Thus, when combined with the circulation shock (and the economic slowdown), the reduction in α may lower the spread that the existing borrowers have to pay out of their buffer capital. However, the reduction in α directly reduces the flow of investors who can purchase liquidated assets in the secondary market. I find that this direct effect is much stronger than the repo spread effect, exacerbating the instability.

Lastly, I consider a reduction in the riskless rate ρ. A reduction in ρ can improve the stability of the economy. It lowers the repo interest that existing borrowers have to pay without affecting the entry to the investor type (I.31). A sufficiently big reduction in ρ can cancel out the impact of the negative circulation shock, and restore the stability of the economy. It further makes all FIs more patient, or equivalently makes time pass faster. Thus the reduction in ρ can also cancel out the impact of the negative shock to either β or α. With the same logic, an increase in the riskless rate ρ works the opposite direction, exacerbating the instability of the economy. A low discount rate ρ can foster future fragility of the economy. With a low ρ, investors choose a low level of buffer capital and take high leverage, therefore making the economy susceptible to repo runs when negative shocks hit the economy.

Positive Feedback Effect This model presents a novel mechanism, where a positive feedback effect escalates haircuts and repo spreads. The intuition is as follows and illustrated in I.8.

A negative circulation shock ($q \downarrow$) instantly increases repo spreads as compensation for the increased opportunity cost of lending. The higher repo spread pressures existing borrowers and pushes them to default earlier. Moreover, it discourages new borrowers from taking high leverage as shown in comparative statics in 3.

For a small shock (i.e., the difference between the degree of circulation before and after the shock is small), the increased supply and decreased demand in the secondary market are absorbed by attracting more borrowers from the primary market. In a stationary equilibrium, aggregate demand is always higher than supply in the secondary market, guaranteeing a positive measure of buyers in the primary market. However, if the shock is sufficiently large, so the degree of circulation drops below a threshold $\bar{q}$, there will be an excess supply of assets that cannot be absorbed by new borrowers.

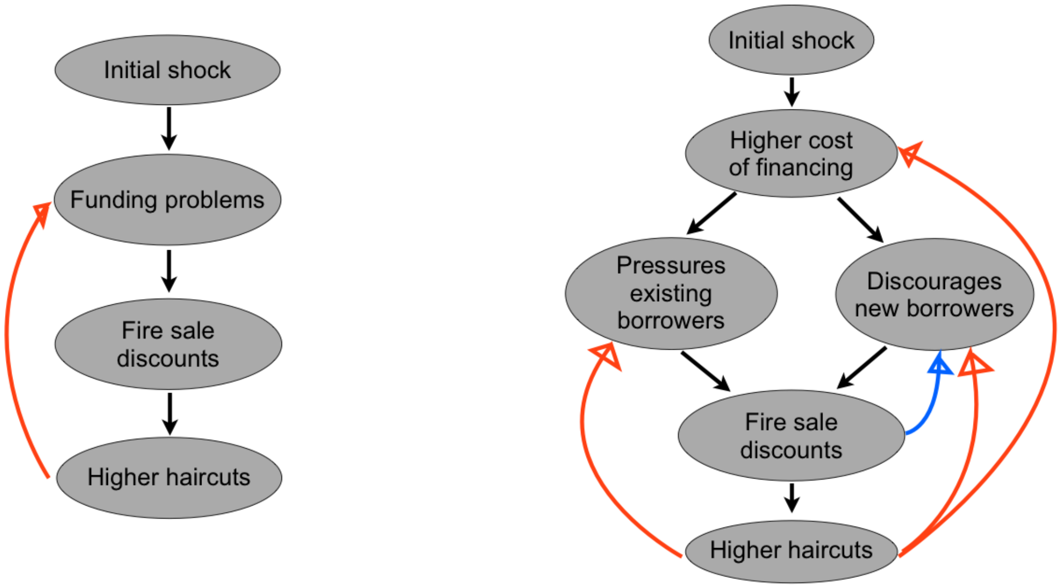


Figure I.8: Positive feedback mechanism through the haircut á la Brunnermeier and Pedersen (2009) (left) and through both the haircut and repo spread (right)

For shocks that cause *any* positive excess supply of liquidated assets, the secondary market price p should adjust downward to clear the market, resulting in fire-sale discounts (or cash-in-the-market pricing). As repo investors can continuously adjust the terms of repos, they adjust the haircut upward one-for-one to protect themselves from the potential cost of immediate liquidation. The higher haircut has three effects: (1) it requires existing borrowers to pay extra cash for the haircut out of buffer capital and pushes them to even earlier bankruptcy, (2) it lowers the extent to which new borrowers can take leverage for a given level of haircut capital, but incentivizes them to increase the amount of haircut capital so that they can take advantage of the present discounts, and (3) it pushes the repo spread even higher, as repo investors understand the fire-sale discount and its profitability, and require compensation for the increased opportunity cost.

The further increase in the repo spread again pressures existing borrowers and discourages new borrowers from taking leverage. Therefore, the escalating repo spread and haircut positively reinforce each other. This positive feedback effect causes an arbitrarily small (but positive) amount of excess supply to drive the haircut up to 1, pushing all existing borrowers into bankruptcy and not allowing new borrowers any leverage. The positive spread is consistent with empirical evidence documented during the recent financial crisis. Gorton and Metrick (2011) document that repo spreads increased from 6.03 bp in the first half of 2007 to 84.27 bp in the second half of 2007, and to 248.29 bp in 2008. Hordahl and King (2008) and Smith (2012) also document the repo spread being mostly positive and sharply increasing during the crisis.

This feedback mechanism results in repo runs where all existing borrowers are forced into bankruptcy. This channel is distinct from the previous literature on financial fragility. Since Diamond and Dybvig (1983), the key component of fragility (or "runs") has been the the lack of sophistication of lenders. More recently, Brunnermeier and Pedersen (2009) show that margins, similar to haircuts, are destabilizing if and only if the lenders are unsophisticated. The general dialogue of the recent financial crisis therefore became that of unsophisticated and panicking lenders. In contrast, I am able obtain fragility when all agents are sophisticated. The key difference between the two models is that the spread is endogenous in my model, which plays a key role in repo runs.

Whether a repo spread should be endogenous or fixed at zero depends on the specific markets considered. In markets where collateral is simple and liquid, such as Treasury securities and stocks, the zero spread assumption may be suitable as the pool of potential lenders is large. However, when financing complex and illiquid collateral such as subprime mortgage-backed-securities, the pool of lenders

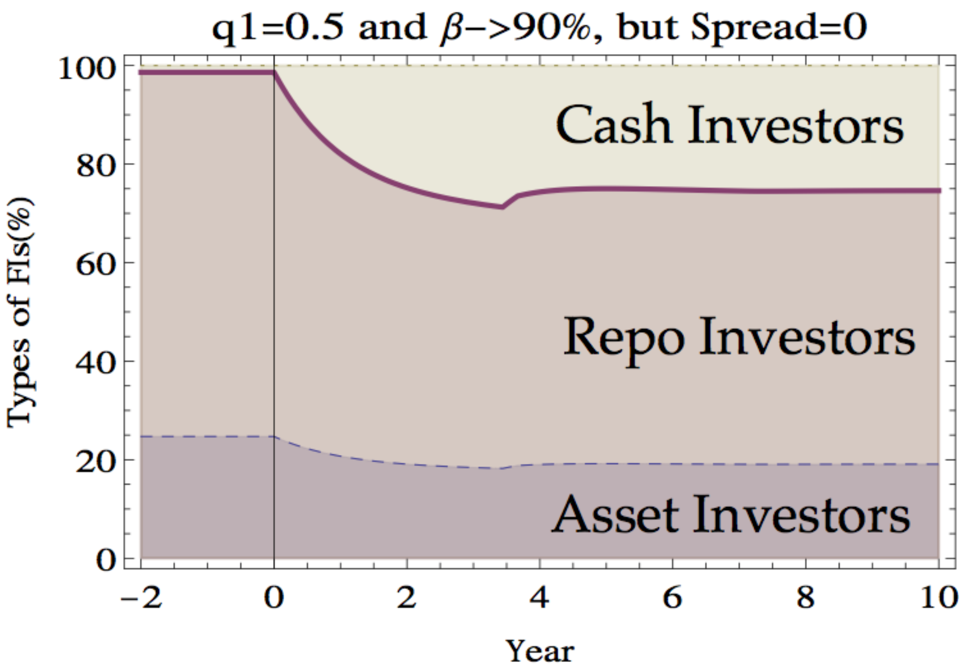


Figure I.9: Transition Path with the Repo Spread Fixed at Zero

who would accept them as collateral is limited and should be taken into account. Therefore the cost of financing *illiquid* collateral should be endogenous and may well be higher than the prevailing riskless rate, resulting in a positive spread.

In I.9, I report the transition path when the repo spread is exogenously fixed at zero to emphasize the role of the endogenous spread in the fragility of the repo market to the circulation shock. Even though the economy suffers both the circulation shock ($q1 = 0.5$) and the economic slowdown (β decreases by 10%), the economy exhibits a stable path.

4.3 Financial Crisis of 2008

Here I briefly describe real world collateral circulation and map the model into the financial crisis of 2008. For additional sources of institutional details, see Fleming and Garbade (2003), Singh and Aitken (2009, 2010), Singh (2011), Copeland, Martin, and Walker (2011) and Adrian, Begalle, Copeland, and Martin (2012).

Collateral circulation is not unique to repo contracts. Financial institutions may receive collateral as a result of other leveraged investment strategies. (In the US, Rule 15c3–3 of the Securities Exchange Act of 1934 limits broker-dealers from circulating their customers' collateral to a certain extent.) The flip side of repo contracts is securities lending, where investors borrow securities from the owners for various purposes (e.g. shorting). In any case, lenders and borrowers may bilaterally agree that lenders are allowed to circulate collateral. Instead of trying to model various strategies, I focus on repos as the repo market is very large and one of the key funding sources for shadow banks.

There are two distinct types of repos: tri-party and bilateral repos. In tri-party repos, a clearing bank provides clearing and settlement services to lenders and borrowers. Typically, lenders are cash rich investors such as money market mutual funds and securities lenders, and borrowers are broker-dealers. Most tri-party repos are backed by highly liquid securities such as Treasury and Agency securities. Lenders accept illiquid collateral from a few large institutions, but haircuts are not adjusted according to market conditions. Moreover, collateral does not circulate in tri-party repos. Once collateral is pledged in tri-party repos, it is not further repledged. Essentially, tri-party repos resemble traditional demand deposits without explicit deposit insurance.

It is bilateral repos that my model is designed to address. In bilateral repos, various financial institutions borrow from and lend to each other and circulate collateral. Many types of collateral are accepted including risky, complex and illiquid securities. They are free to adjust the terms of repos according to market conditions. Note that a broker-dealer may receive collateral from a hedge fund and

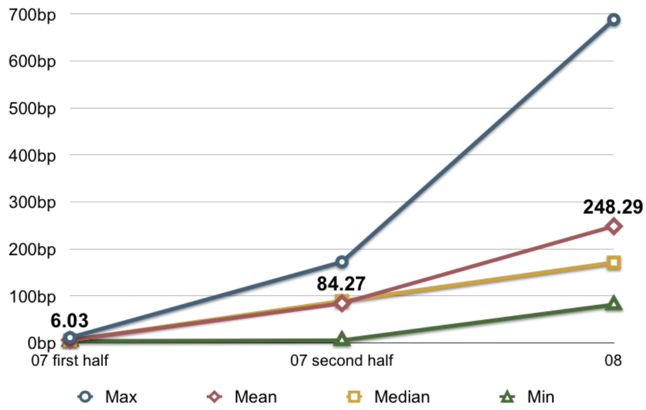


Figure I.10: Repo Spread during the Financial Crisis. Source: Gorton and Metrick (2011).

pledge it in tri-party repos, making tri-party repos the end-point of a collateral chain.

The differences between the two repos are crucial in understanding what happened in the repo markets during the recent crisis. A sharp contraction in bilateral repos has been documented by Gorton and Metrick (2010a,b, 2011, 2012). The average haircut increased from 0 to 45% and certain classes of securities stopped being used as collateral entirely. The average repo spread increased from under 10 bp to over 200 bp, with a maximum close to 700 bp as in I.10.[20] However, Krishnamurthy, Nagel, and Orlov (2012) find that tri-party repos remained largely stable, consistent with the fact that most tri-party repos were backed by safe collateral that became more valuable during the crisis. On the other hand, funding for illiquid collateral completely dried up, but only a tiny fraction of illiquid collateral was financed by tri-party repos to begin with. They conclude the problem in the repo market was more like a credit crunch caused by dealer banks tightening their funding for the borrowers than a traditional bank run caused by panicked depositors.

[20]See also Adrian and Shin (2010).

My model is consistent with their findings that bilateral repos with collateral circulation can be fragile, and perhaps more fragile than tri-party repos, with collateral circulation exacerbating the leverage cycle. This model suggests a mechanism for such a credit crunch among shadow banks in conjunction with the fact that collateral circulation suffered a large, negative shock during the crisis. Repo runs during the financial crisis of 2008 may not have been caused by panicking and uninformed lenders, but rather by the inherent fragility of collateral circulation.

5 Welfare

When a repo run occurs after a shock to collateral circulation, all existing repo investors as well as borrowers are forced into bankruptcy since collateral becomes worthless. If all the existing repo investors could get together, they would prefer to avoid a run by not raising the repo spread. The fact that repo investors can update the terms of the repo instantly and there is a continuum of repo investors who act competitively is the source of constrained inefficiency. Of course, given the shock to collateral circulation and changes in market prices, the optimal response of repo investors is to make their repos safe and get their required compensation. Collectively, however, this creates a repo run that forces the existing repo investors into defaults upon the shock. That is, a repo run equilibrium is constrained inefficient.[21]

Proposition 5. *A repo run equilibrium is constrained inefficient. Upon the arrival of a shock, a social planner, who cannot directly affect collateral circulation,*

[21] Relatedly, Lorenzoni (2008) and Korinek (2011) show such pecuniary externalities (a type of externalities that operate through prices rather than real resources) can cause real inefficiency when agents have financial constraints.

can obtain welfare gains by taxing existing repo investors to support the interest payment of the existing borrowers.

The fragility of collateral circulation and the inefficiency of a repo run perhaps echo Friedman (1960), who strongly argued against the private creation of money. He argued that allowing private provision of circulating liabilities generates indeterminacy of equilibrium and excess volatility. Thus the creation of money should be segregated from all private market activity and solely left to the government. On the other hand, Hayek (1976) and Fama (1980) argued that even the creation of money can be done in markets efficiently.[22]

Collateral circulation, effectively creating money within the system of financial institutions, does create extreme fragility to a sudden shock as Friedman argued. However, banning all institutions from bilaterally agreeing to circulate collateral can be harmful. Collateral circulation in normal times creates liquidity in the system and real multiplier effects, allowing more investment in the economy, as Hayek and Fama predicted. That is, there is a trade-off between the economic growth and financial stability. Frictionless collateral circulation leads to more investment and promotes growth, but increases the likelihood that the economy will be exposed to an inefficient repo run.

5.1 Policy Implications

Since the recent crisis academics and policy makers have focused on the level of leverage and capital requirements for financial institutions. The repo market's behavior during the crisis, the steep decline in collateral for circulation as docu-

[22]See also Azariadis, Bullard, and Smith (2001).

mented here, and the fragility the model produces all point to the importance of collateral circulation and its fragility. Several broker-dealers play an important role in collateral circulation, and circulating collateral is more important for some institutions than others. Thus it may be necessary for regulators to keep track of collateral circulation with a special focus on financial institutions that are critical in collateral circulation and heavily rely on it. In other words, these institutions are systemically important financial institutions (or SIFIs.)

The result on the tradeoff between growth and stability implies that simply banning collateral circulation is not ideal and can be very costly. Rather, it suggests there is a role for the lender of last resort. The central bank can provide funding in a crisis to avoid runs while supporting economic growth in good times. Suppose in a crisis collateral circulation suddenly contracts. For example, borrowers become worried about getting their collateral back and do not allow lenders to circulate it, lenders become cautious and sit on collateral, or critical institutions in the network fail. The central bank can restore collateral circulation by lending against collateral to jittery borrowers, lenders who wish to circulate their collateral, and the institutions whose prime brokers are failing. In other words, the central bank becomes a part of the system temporarily.

In fact, this is what the Federal Reserve did during the recent crisis. Emergency lending facilities such as the Primary Dealer Credit Facility (PDCF) and the Term Securities Lending Facility (TSLF) were set up to inject liquidity into the system through primary dealers.[23] The programs lent cash or Treasury securities against a range of collateral. Rapid policy responses may be essential during a crisis, so it

[23]Primary dealers are twenty or so financial institutions that frequently trade with the Federal Reserve system in its implementation of monetary policy.

may be preferable to have such funding programs permanently on standby rather than as emergency facilities. The terms for funding - both the haircut and repo rate - can be set so that they are only favorable in a crisis, similar to discount windows. However, the existence of such programs, unlike deposit insurance, is not sufficient to stop runs from happening since repo runs are not driven by mere perceptions. The programs should be ready into inject liquidity to the system.

The idea that the lender of last resort should lend against collateral goes back to Bagehot (1878) (see also Thornton (1802) and Goodhart (1999)). He argued that to avert panic, central banks should lend to solvent firms against good collateral early, without limit, and at high rates. In contrast, my model suggests that to stop a run from happening after a shock to collateral circulation, the lender of last resort should lend (1) against a variety of collateral, including risky, complex and illiquid securities, (2) at low rates, perhaps even zero spread, but (3) with an appropriate haircut to make the loan safe. From the divergence across repo markets during the crisis, the strain appears to be concentrated among illiquid securities. Even loans against risky collateral can be safe as long as the haircut is set accordingly. Financial institutions should be allowed to pledge a portfolio of securities, thus lowering the haircut required to make the loan safe. Central banks have an advantage as a lender, compared to those in the system, in that they have flexibility in their horizon for liquidating collateral rather than having to liquidate collateral immediately.

The funding programs could limit the range of acceptable collateral to those that may be risky and illiquid but have high hold-to-maturity value, exploiting the advantage of their flexibility. The liquidity injected could be trickled down to more illiquid assets, as financial institutions that own such collateral can use liquidity to purchase other assets at deep discounts. This will be especially profitable for the institutions in a crisis since the loans have low rates. Without any programs, collateral circulation poses the problem of too-many-to-fail (Acharya and Yorulmazer (2007)), where numerous financial institutions, regardless of their size, fail together and central banks end up supporting failing institutions ex-post. Thus, having a funding program with eligible collateral being specified in advance would reward financial institutions for staying solvent and owning sound collateral in a crisis, mitigating moral hazard problems.

6 Conclusion

In this paper I have studied the economic importance of collateral circulation in a dynamic competitive economy. A continuum of ex-ante symmetric financial institutions borrow from and lend to one another against illiquid collateral. As in Kiyotaki and Moore (2001), the institutions receive random arrivals of investment opportunities, which allow them to invest in risky, profitable, and illiquid assets. That institutions need an opportunity to make investments is related to the slow-moving capital literature following Mitchell, Pedersen, and Pulvino (2007) and Duffie (2010). Due to financing constraints that limit the ability of institutions to issue equity or risky debt, institutions have an incentive to conserve liquidity (cash) for future investments. When repos are liquid (or have high "moneyness") institutions substitute cash with repos. The capital invested in repos allows other institutions to take leverage and scale up their investments. While collateral circulation increases steady state investment, it can make the financial system fragile because less capital is sitting on the sidelines waiting for investment opportunities to arrive. The positive feedback loop between the repo spread and fire-sale discounts can result in inefficient repo runs, suggesting a role for a lender of the last resort.

This paper leaves several important questions for future research. One such question is what determines repo liquidity, or the ease of collateral circulation. Many real-world frictions, such as search frictions, network effects, and imperfect competition, can potentially play a role. Given the inefficiency in a competitive equilibrium, it is also an open question whether large institutions will internalize fire-sale externalities, possibily improving welfare. Another question is what happens when there are more than one class of assets. Financial institutions in reality own and borrow against a portfolio of assets, and shocks to collateral circulation may limit the extent to which they can do cross margining. This can create contagions across assets and institutions, similar to Kyle and Xiong (2001). Finally, it is important to consider what happens when shocks are anticipated. Ongoing work suggests when shocks are sufficiently infrequent, the results are largely unchanged, although the financial system is slightly more resilient.

A Appendix

A.1 Proofs

Proof of Lemma 1. For simplicity of the exposition, denote

$$\bar{s} = \left(\frac{1}{h} - 1\right)(\rho + s) \tag{37}$$

Then we can write the leveraged value function as

$$V_{leveraged} = 1 + \left[\left(\frac{\beta\bar{R} - \bar{s}}{\rho + \beta}\right)\left(1 - e^{-(\rho+\beta)T_\delta}\right) - 1\right]x$$

If $\frac{\beta\bar{R}-\bar{s}}{\rho+\beta} \leq 1$, we have $x = 0$ and $V_{leveraged} = 1$. Otherwise, the F.O.C. with respect to x is

$$\left(\frac{\beta\bar{R} - \bar{s}}{\rho + \beta}\right)\left(1 - e^{-(\rho+\beta)T_\delta}\right) - 1 = -T_\delta^{'}x\left(\beta\bar{R} - \bar{s}\right)e^{-(\rho+\beta)T_\delta}$$

Again from the solvency constraint, we have $-xT_\delta^{'} = \frac{1}{x\bar{s}e^{-\rho T_\delta}}$, so the f.o.c. is now

$$x\left[\left(\frac{\beta\bar{R} - \bar{s}}{\rho + \beta}\right)\left(1 - e^{-(\rho+\beta)T_\delta}\right) - 1\right] = \left(\beta\frac{\bar{R}}{\bar{s}} - 1\right)e^{-\beta T_\delta} \tag{38}$$

We can see that the R.H.S. of the f.o.c. is the same as the value function $V_{levered}$. Thus, we can rewrite the value function as

$$V_{levered} = 1 + e^{-\beta T_\delta^*}\left(\beta\frac{\bar{R}}{\bar{s}} - 1\right) = 1 + \left(\beta\frac{\bar{R}}{\bar{s}} - 1\right)\theta^* = 1 + \frac{\beta}{\rho + s}\left(\frac{R}{1-h} - 1\right)\theta^*$$

Substitute $x = \left(\frac{1-e^{-\rho T_\delta}}{\rho}\bar{s} + 1\right)^{-1}$ and $-xT_\delta^{'} = \frac{1}{x\bar{s}e^{-\rho T_\delta}}$ into (38) and then $\theta^*(\rho, \beta, s, h, R) \equiv$

$e^{-\beta T_{\delta}^{*}} \in (0,1)$ solves the following equation:

$$f(\theta) \equiv \left(\frac{1}{\rho}+\frac{1}{\bar{s}}\right)\theta - \left(\frac{1}{\rho}-\frac{1}{\rho+\beta}\right)\theta^{1+\frac{\rho}{\beta}} = \frac{1}{\rho+\beta} - \frac{1}{\beta\bar{R}-\bar{s}} \tag{39}$$

Here, $f(\theta)$ is positive, increasing for $\forall\theta \in (0,1)$ and

$$\{f(\theta) \mid \theta \in (0,1)\} = \left(0, \frac{1}{\rho+\beta}+\frac{1}{\bar{s}}\right)$$

Moreover, since $\frac{\beta\bar{R}-\bar{s}}{\rho+\beta} > 1$, the R.H.S. of (39)

$$\frac{1}{\rho+\beta} - \frac{1}{\beta\bar{R}-\bar{s}} \in \left(0, \frac{1}{\rho+\beta}+\frac{1}{\bar{s}}\right)$$

Therefore, there exists a unique solution $\theta^{*} = e^{-\beta T_{\delta}^{*}} \in (0,1)$ that solves (39).

The value of leveraged investment $V_{leveraged}$ is decreasing in s and strictly decreasing when $V_{leveraged} > 1$. This is because when the repo spread lowers, the borrower can continue to choose the same amount of haircut capital and achieve the same expected payoff. Moreover, he can consume the difference of the repo spreads. That is,

$$\frac{dV_{leveraged}}{ds} \begin{cases} < 0 & \text{if } s < \frac{\beta(R-1)-\rho}{1-h} \\ = 0 & \text{otherwise} \end{cases}$$

Thus we have (I.12).

Lastly, that the investor's threshold $\hat{s} > 0$ is equivalent to

$$\left[\frac{\beta\bar{R}}{\left(\frac{1}{h}-1\right)\rho} - 1\right] \cdot \theta^{*}(0) > \left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}} - 1$$

That is, from (39),

$$\left(\frac{1}{\rho}+\frac{1}{\left(\frac{1}{h}-1\right)\rho}-\frac{\beta}{\rho\left(\rho+\beta\right)}\left[\frac{\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}}-1}{\frac{\beta\bar{R}}{\left(\frac{1}{h}-1\right)\rho}-1}\right]^{\frac{\rho}{\beta}}\right)\left[\frac{\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}}-1}{\frac{\beta\bar{R}}{\left(\frac{1}{h}-1\right)\rho}-1}\right]<\frac{1}{\rho+\beta}-\frac{1}{\beta\bar{R}-\left(\frac{1}{h}-1\right)\rho}$$

After algebra, we get

$$0<\frac{\beta}{\rho\left(\rho+\beta\right)}\left[\frac{\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}}-1}{\frac{\beta\bar{R}}{\left(\frac{1}{h}-1\right)\rho}-1}\right]^{\frac{\rho}{\beta}}\iff\beta\frac{R}{\bar{p}}>\left(1-h\right)\left(\rho+\beta\right)$$

which is true for $\forall h\in[0,1)$ by the assumption that assets are profitable. Therefore, $\hat{s}>0$. □

Proof of Lemma 2. From $V_L=V_W$, (I.16) and (I.14), we have

$$\check{s}\left(V_I\right)=\left(1-q\right)\frac{\alpha\rho}{\alpha+\rho}\left(V_I-1\right)$$

Moreover, the solution to the fixed point problem such that

$$\check{s}=\left(1-q\right)\frac{\alpha\rho}{\alpha+\rho}\left(V_I\left(\check{s}\right)-1\right)$$

always exists and unique because the L.H.S. is strictly increasing in s and the R.H.S. is nonnegative and weakly decreasing in s by Lemma 1. □

Proof of Proposition 1. It is straightforward. The unique of equilibrium follows the uniqueness of solution to the first order condition from Lemma 1. □

Proof of Proposition 2. (i) $\frac{ds^*}{dq}\leq 0$ is straightforward. From the Implicit Function

Theorem and 61, it is enough to show $\frac{\partial V_B}{\partial s} \leq 0$. Suppose not. That is, $s_1 > s_2$ and

$$V_B(s_1) = V_B(x^*(s_1), T^*_\delta(s_1)) > V_B(s_2) = V_B(x^*(s_2), T^*_\delta(s_2))$$

Then at the lower repo rate s_2, borrower can still choose $x^*(s_1)$ and $T^*_\delta(s_1)$ and consume the remaining capital $\epsilon > 0$ and get

$$V_B(s_2) = V_B(s_1) + \epsilon > V_B(s_1)$$

It is contradiction.

(ii) Now consider θ^*. Recall the F.O.C.

$$\left(\frac{1}{\rho} + \frac{1}{\bar{s}^*}\right)\theta^* - \left(\frac{1}{\rho} - \frac{1}{\rho+\beta}\right)\theta^{*\left(1+\frac{\rho}{\beta}\right)} = \frac{1}{\rho+\beta} - \frac{1}{\beta\bar{R} - \bar{s}^*}$$

Thus we have

$$\theta' \equiv \frac{d\theta^*}{d\bar{s}^*} = \frac{\frac{\theta^*}{\bar{s}^{*2}} - \frac{1}{\left(\beta\bar{R} - \bar{s}^*\right)^2}}{\frac{1}{\bar{s}^*} + \frac{1}{\rho} - \frac{\theta^{*\frac{\rho}{\beta}}}{\rho}} = \frac{\theta^* - \left(\frac{\beta\bar{R}}{\bar{s}} - 1\right)^{-2}}{\bar{s} + \frac{\bar{s}^2}{\rho}\left(1 - \theta^{*\frac{\rho}{\beta}}\right)} \tag{40}$$

$\theta^{'} > 0$ if and only if $\theta^* > \left(\frac{\beta\bar{R}}{\bar{s}^*} - 1\right)^{-2}$ (and $\theta^{''} < 0$ for $\forall \bar{s} < \beta\bar{R}$). That is, $V_{levered}(\bar{s}^*) - 1 = \left(\frac{\beta\bar{R}}{\bar{s}^*} - 1\right)\theta^* > \left(\frac{\beta\bar{R}}{\bar{s}^*} - 1\right)^{-1} = -1 + \frac{\beta\bar{R}}{\beta\bar{R} - \bar{s}^*}$ or equivalently,

$$V_{leveraged}(\tilde{s}^*) > \frac{\beta\bar{R}}{\beta\bar{R} - \bar{s}^*}$$

Since the L.H.S. is decreasing in $\bar{s}$ and the R.H.S. is increasing in $\bar{s}$ (for $\bar{s} < \beta\bar{R}$), there exists $s_\theta > 0$ such that

$$\theta' \begin{cases} \geq 0 & \text{if } \bar{s} \leq s_\theta \\ \\ < 0 & \text{otherwise} \end{cases}$$

where $s_\theta > 0$ solves

$$\theta^*(s_\theta) = \left(\frac{\beta \bar{R}}{\tilde{s}_\theta} - 1\right)^{-2}$$

In a type 1 equilibrium, all equilibrium repo spreads $s < s_b = \min\left\{\hat{s}, \hat{\hat{s}}\right\}$ where $\hat{s}$ and $\hat{\hat{s}}$ are defined as I.13 and I.18. Thus, it is enough to show that $s_\theta > \hat{s}$. Define

$$\theta_1(s) \equiv \left(\frac{\beta \bar{R}}{\left(\frac{1}{h} - 1\right)(\rho + s)} - 1\right)^{-2} = \left(\frac{\beta (R - 1 + h)}{(1 - h)(\rho + s)} - 1\right)^{-2}$$

and

$$\theta_2(s) \equiv \frac{s}{\rho}\left(\frac{\beta (R - 1 + h)}{(1 - h)(\rho + s)} - 1\right)^{-1}$$

Then s_θ is where $\theta_1(s)$ and $\theta^*(s)$ meet and $\hat{s}$ is where $\theta_2(s)$ and $\theta^*(s)$ meet. Define

$$s_{12}(h, \rho, \beta, R) \equiv \frac{\beta (R - 1 + h)}{2(1 - h)}\left[1 - \sqrt{1 - \frac{4\rho(1 - h)}{\beta(R - 1 + h)}}\right] - \rho = 2\rho\left[\frac{1 - \sqrt{1 - Y}}{Y}\right] - \rho \tag{41}$$

which is the solution to $\theta_1(s_{12}) = \theta_2(s_{12})$ where

$$Y \equiv \frac{4\rho(1 - h)}{\beta(R - 1 + h)} \text{ or } \frac{1}{1 - h} = \frac{1}{R}\left(1 + \frac{4\rho}{\beta Y}\right) \tag{42}$$

and $Y \in \left(0, \frac{4\rho}{\beta(R-1)}\right)$ is decreasing in h (Note that $Y < 1$ from $R - 1 \geq 4\frac{\rho}{\beta}$.) Then from $\theta^*(0) > \theta_1(0) > \theta_2(0) = 0$, $s_\theta > \hat{s}$ is equivalent to $s_{12} < s_\theta$. Since $\theta' > 0$ for

$s < s_\theta$, it is enough to show

$$\theta^*(0) > \left(\frac{s_{12}}{\rho}\right)^2 = \left\{2\left[\frac{1-\sqrt{1-Y}}{Y}\right]-1\right\}^2 \tag{43}$$

as $\theta_1(s_{12}) = \theta_2(s_{12}) = \left(\frac{s_{12}}{\rho}\right)^2$. Further define

$$Z \equiv \left\{2\left[\frac{1-\sqrt{1-Y}}{Y}\right]-1\right\}^2 \text{ or } Y = \frac{4\sqrt{Z}}{\left(1+\sqrt{Z}\right)^2}$$

$Z \in \left(0, \bar{Z} \equiv \left\{\frac{\beta(R-1)}{2\rho}\left[1-\sqrt{1-\frac{4\rho}{\beta(R-1)}}\right]-1\right\}^2\right)$ is increasing in Y. Then from the FOC, (43) is equivalent to

$$Z\left[\frac{1}{R}\left(1+\frac{\rho}{\beta}\frac{\left(1+\sqrt{Z}\right)^2}{\sqrt{Z}}\right)-\left(\frac{\beta}{\rho+\beta}\right)Z^{\frac{\rho}{\beta}}\right] < \frac{\rho}{\rho+\beta}\left[1+\frac{(\rho+\beta)h}{\beta(R-1)-\rho}\right] \tag{44}$$

Claim 1. The LHS in (44) is increasing in Z and the inequality holds as $Z \to \bar{Z}$. Thus the inequality holds for $\forall Z$.

Proof of Claim1. First show that the LHS is increasing in Z. Consider the derivative of the LHS with respect to Z.

$$\frac{1+\frac{\rho}{\beta}\left(\frac{3\sqrt{Z}+\frac{1}{\sqrt{Z}}+4}{2}\right)}{1+\frac{\rho}{\beta}r} - Z^{\rho/\beta} \geq 0 \tag{45}$$

where we write $r = \beta(R-1)/\rho$ and $\bar{Z}(r) \equiv \left\{\frac{r}{2}\left[1-\sqrt{1-\frac{4}{r}}\right]-1\right\}^2$. It is increasing in $\frac{\rho}{\beta}$, so the inequality holds if and only if $\frac{\rho}{\beta}$ is greater than some threshold.

Moreover,

$$\lim_{\rho/\beta\to 0}\frac{1+\frac{\rho}{\beta}\left(\frac{3\sqrt{Z(r)}+\frac{1}{\sqrt{Z(r)}}+4}{2}\right)}{1+\frac{\rho}{\beta}r} = 1 \geq \lim_{\rho/\beta\to 0} Z(r)^{\rho/\beta} = 1$$

Thus, the inequality (45) holds for $\forall Z$. Finally show that the inequality in (44) holds as $Z \to \bar{Z}$. As Z is increasing in Y and Y is decreasing in h, Z is decreasing in h. Thus, we only need to show

$$\lim_{h\to 0}\left(\frac{s_{12}}{\rho}\right)^2\left[\frac{1}{1-h}-\left(\frac{\beta}{\rho+\beta}\right)\left(\frac{s_{12}}{\rho}\right)^{2\frac{\rho}{\beta}}\right]-\frac{\rho}{\rho+\beta}\left[1+\frac{(\rho+\beta)\,h}{\beta\,(R-1)-\rho}\right]=-\frac{\rho}{\rho+\beta}<0$$

which follows $\lim_{h\to 0}\left(\frac{s_{12}}{\rho}\right)^2=1$.

(iii) Now consider x^*. From

$$V_{leveraged}=1+\left[\left(\frac{\beta\bar{R}-\bar{s}}{\rho+\beta}\right)\left(1-\theta^{\frac{\rho+\beta}{\beta}}\right)-1\right]x$$

We know

$$\frac{dV_{leveraged}}{ds}=\left[\left(\frac{\beta\bar{R}-\bar{s}}{\rho+\beta}\right)\left(1-\theta^{\frac{\rho+\beta}{\beta}}\right)-1\right]\frac{dx}{ds}+\left[d\left[\left(\frac{\beta\bar{R}-\bar{s}}{\rho+\beta}\right)\left(1-\theta^{\frac{\rho+\beta}{\beta}}\right)\right]-1\right]x<0$$

Given $\frac{d\theta}{ds}>0$, we have $\left[d\left[\left(\frac{\beta\bar{R}-\bar{s}}{\rho+\beta}\right)\left(1-\theta^{\frac{\rho+\beta}{\beta}}\right)\right]-1\right]x>0$. Thus, $\frac{dx}{ds}<0$ follows.

(iv) $\frac{d\mu_I}{dq}$ Here

$$\mu_I=\left\{\frac{\beta}{\alpha\,(1-\theta)}+1+(1-q)\left(\frac{1}{h}-1\right)x\right\}^{-1}$$

Then we have $\frac{d\mu_I}{dq}=\frac{\partial\mu_I}{\partial q}+\frac{dx}{dq}\frac{\partial\mu_I}{\partial x}+\frac{d\theta}{dq}\frac{\partial\mu_I}{\partial\theta}$. Since

$$\frac{\partial\mu_I}{\partial q}=\left(\frac{1}{h}-1\right)x\mu_I^2,\ \frac{\partial\mu_I}{\partial x}=-(1-q)\left(\frac{1}{h}-1\right)\mu_I^2\text{ and }\frac{\partial\mu_I}{\partial\theta}=-\frac{\beta}{\alpha\,(1-\theta)^2}\mu_I^2$$

Thus, $\frac{d\mu_I}{dq} > 0$ is equivalent to

$$\left(\frac{1}{h}-1\right)\left(x-(1-q)\frac{dx}{dq}\right) > \frac{\beta}{\alpha(1-\theta)^2}\frac{d\theta}{dq}$$

Since $\frac{d\theta}{dq} < 0$, it is enough to show $x \geq \frac{dx}{dq}$ That is, $x(q) \leq e^q$. We know $x(q) \leq 1$ and $e^q \geq 1$ for $\forall q$.

$\square$

Proof of Proposition 3. On the stable path for $\forall t \in \left[T, \hat{T}\right]$, the supply of the secondary market is given by $\alpha_0 x_0^* \left(\frac{1}{h}-1\right)(\mu_C^* + q_0\mu_R^*)\Theta^*(t)$ and the demand for asset is given by $\alpha_1 x_1^*(\mu_C(t) + q\mu_R(t))$. Thus the necessary condition for the stable path I.35 follows.

$\square$

Proof of Proposition 4. To show this is an equilibrium path, I only need to show (i) With $\lim_{t\to 0+} h(t)^* = 1$, the borrowers cannot take leverage anyway. Therefore, they (at least weakly) prefer $\lim_{t\to 0+} x_t^* = 0$ to any other alternatives, hence it is optimal.

(ii) The demand for the secondary market is as below (the buyers of the secondary market still need to incur the transaction cost, making the effective price h_0, which is not zero)

$$\alpha \cdot \lim_{t\to 0+}(\mu_C(t) + q\mu_R(t)) \cdot \frac{1}{h_0}dt = \frac{\alpha}{h_0}dt$$

and the supply is

$$\int_{-T_\delta^-}^{0} e^{\beta\tau}\alpha_0(\mu_{C0} + q_0\mu_{R0}) \cdot \frac{x_0}{h_0}d\tau$$

$$= \alpha_0 (\mu_{C0} + q_0 \mu_{R0}) \cdot \frac{x_0}{h_0} \int_{-T_{\delta 0}}^{0} e^{\beta\tau} d\tau = \alpha_0 (\mu_{C0} + q_0 \mu_{R0}) \cdot \frac{x_0}{h_0} \left(\frac{1 - e^{-\beta T_\delta^-}}{\beta} \right)$$

There is an excess supply and the price reached at the lower bound.

(iii) From (61), the repo rate at which repo investors and cash investors are indifferent is

$$\lim_{t \to 0+} s_t^* = \rho + \frac{\alpha\rho}{\alpha + \rho} (1 - q) \left[\left(\frac{\beta}{\rho + \beta} \right) \frac{R}{h_0} - 1 \right]$$

As there is always a positive measure of cash investors, the equilibrium repo rate is determined so as to repo investors and cash investors are indifferent.

□

Proof of Proposition 6. From Lemma 3, $\mu_C (s_l) > 0$ if and only if

$$q (1 - \theta^* (s_l)) x^* (s_l) \leq \frac{\beta}{\alpha \left(\frac{1}{h} - 1 \right)} = \frac{\beta}{\alpha \left(\frac{1}{\phi} - 1 \right)}$$

where the last equality follows the equilibrium property $h = \phi$ (I.26). Define

$$k \equiv \max_{\alpha, \beta, \phi, s} (1 - \theta^*) x^* \in [0, 1)$$

Then the sufficient condition for $\mu_W (s_l) > 0$ is that

$$\phi \geq \hat{\phi} (\alpha, \beta) \equiv \left(\frac{\beta}{\alpha k} + 1 \right)^{-1}$$

Now we want to find the condition under which $s_l < s_b \equiv \min \left\{ \hat{s}, \hat{\hat{s}} \right\}$. There are two cases: (i) $\hat{s} \leq \hat{\hat{s}}$ and $\hat{s} = s_b$. Then the condition is equivalent to $s_l < \hat{s} \leq \hat{\hat{s}}$ and (ii) $\hat{\hat{s}} \leq \hat{s}$ and $\hat{\hat{s}} = s_b$. Then since V_I is decreasing in s, (I.17) and (I.18) imply that

$\hat{\hat{s}} \geq \check{s}$. Thus, $\check{s} < s_b$ is equivalent to $s_l < \hat{s} \leq \hat{\hat{s}}$.

If $q = 1$, $s_l = 0$ and $s_b > 0$, which implies that $s_l < s_b$. Now for $q < 1$, the sufficient condition is that $s_l < s_b$, which is equivalent to $s_l\left(V_{unlevered}\right) < \hat{s}$. That is,

$$V_{leveraged}\left(s_l\right) > \left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}}$$

where

$$s_l = (1-q)\left(\frac{\alpha\rho}{\rho+\alpha}\right)\left(\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}} - 1\right)$$

From (I.12), denoting $\check{V}_l \equiv V_{leveraged}\left(s_l\right)$, we have

$$\check{V}_l - 1 = \left(\frac{\beta\bar{R}}{\left(\frac{1}{h}-1\right)\rho\left[1+(1-q)\left(\frac{\alpha}{\rho+\alpha}\right)\left(\check{V}_l - 1\right)\right]} - 1\right)\theta^*\left(s_l\right) > \left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}} - 1$$

thus,

$$\theta^*\left(s_l\right) > \frac{\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}} - 1}{\frac{\beta\bar{R}}{\left(\frac{1}{h}-1\right)\rho\left[1+(1-q)\left(\frac{\alpha}{\rho+\alpha}\right)\left(\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}}-1\right)\right]} - 1}$$

From the F.O.C. of the investor, it is equivalent to

$$\left(\frac{1}{\rho} + \frac{1}{\left(\frac{1}{h}-1\right)\left(\rho+s_l\right)} - \frac{\beta}{\rho\left(\rho+\beta\right)}\left[\frac{\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}} - 1}{\frac{\beta\bar{R}}{\left(\frac{1}{h}-1\right)\left(\rho+s_l\right)} - 1}\right]^{\frac{\rho}{\beta}}\right)\left[\frac{\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}} - 1}{\frac{\beta\bar{R}}{\left(\frac{1}{h}-1\right)\left(\rho+s_l\right)} - 1}\right]$$

$$< \frac{1}{\rho+\beta} - \frac{1}{\beta\bar{R} - \left(\frac{1}{h}-1\right)\left(\rho+s_l\right)}$$

After algebra, we get

$$\left(\frac{s_l}{\rho+s_l}\right)\left[\frac{\beta\left(\frac{R}{\bar{p}} - 1 + h\right)}{(1-h)\left(\rho+s_l\right)} - 1\right]^{\frac{\rho}{\beta}} < \frac{\left[\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}} - 1\right]^{1+\frac{\rho}{\beta}}}{\frac{R}{\bar{p}} - 1}$$

Substitute s_l into the inequality and get

$$\frac{\beta\left(\frac{R}{\bar{p}}-1+h\right)}{(1-h)} < \rho\left\{1+(1-q)\left(\frac{\alpha}{\rho+\alpha}\right)\left[\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}}-1\right]\right\}$$

$$\cdot\left[1+\left[\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}}-1\right]\left\{\frac{\left(\frac{\beta}{\rho+\beta}\right)\frac{R}{\bar{p}}-1}{\frac{R}{\bar{p}}-1}+\frac{1}{(1-q)\left(\frac{\alpha}{\rho+\alpha}\right)\left(\frac{R}{\bar{p}}-1\right)}\right\}^{\frac{\beta}{\rho}}\right]$$

Here, the L.H.S. is increasing in h and the R.H.S. is independent of h. Therefore, there exists $\check{h}$ such that the condition is equivalent to

$$h \leq \check{h} \text{ if and only if } s_l < \hat{s}$$

$\square$

Proof of Proposition 5. Suppose the economy can be described with the parameters before the shock $\{q_0,\alpha_0,\beta_0,\sigma_0,\rho_0,R_0,\phi_0\}$ and after the shock $\{q,\alpha,\beta,\sigma,\rho,R,\phi\}$. By Proposition 1, a unique stationary equilibrium exists in the economy before the shock and is denoted by $\left(s_0^*,x_0^*,h_0^*,p_0^*,\{\mu_{i0}^*\}_{i\in\{I,R,C\}}\right)$. In a repo run defined in Proposition 4, all the existing borrowers and repo investors before the shock go bankrupt and exit the economy. It can improve Pareto efficiency to tax existing repo investors and support existing borrowers so that they can afford to pay s_1^*, which is the repo spread at a stationary equilibrium with the parameters after the shock. As all cash of the repo investors is lent to borrowers, taxing repo investors can be implemented by forcing repo investors to forgive some of the principal and/or to accept lower repo interests than they would require otherwise.

By Proposition 3, we have without intervention

$$\frac{\alpha_1 x_1^*}{\alpha_0 x_0^*}\left(\frac{\mu_C(t)+q\mu_R(t)}{\mu_C^*+q_0\mu_R^*}\right)\geq \exp\left\{-\left[(\beta-\beta_0)(t-T)+\beta_0 T_\delta^*(t_0)\right]\right\} \tag{46}$$

By forcing repo investors to not receive any interests until the investments mature and to not circulate their collateral even when they receive an opportunity, we can increase $T_\delta^*(t_0)\to\infty$ for $\forall t_0<T$ as $s_1^*=0$ in (I.31) and (I.35), meaning all the existing investors will never default. Since the LHS in (46) is always positive, the economy exhibits the stable path. Therefore, with the intervention, all the existing borrowers and repo investors are strictly better off, ex-ante by $E\left\{\int_0^\infty e^{-\rho\tau_\beta}\tilde{R}x_0^* d\tau_\beta\right\}=\left(\frac{\beta}{\rho+\beta}\right)\left(1+\frac{R-1}{h}\right)x_0^*$ for each borrower and $E\left\{\int_0^\infty e^{-\rho\tau_\beta}1d\tau_\beta\right\}=\frac{\beta}{\rho+\beta}$ for each repo investor. □

Proof of Lemma 3. Combining (I.22) - (I.24), I get

$$\mu_C=\frac{\beta-\alpha q(1-\theta^*)\left(\frac{1}{h}-1\right)x^*}{\beta+\alpha(1-\theta^*)\left[1+(1-q)\left(\frac{1}{h}-1\right)x^*\right]} \tag{47}$$

$$\mu_R=\frac{\alpha(1-\theta^*)\left(\frac{1}{h}-1\right)x^*}{\beta+\alpha(1-\theta^*)\left[1+(1-q)\left(\frac{1}{h}-1\right)x^*\right]} \tag{48}$$

and

$$\mu_I=\frac{\alpha(1-\theta^*)}{\beta+\alpha(1-\theta^*)\left[1+(1-q)\left(\frac{1}{h}-1\right)x^*\right]}, \tag{49}$$

if $\beta-\alpha q(1-\theta^*)\left(\frac{1}{h}-1\right)x^*>0$ and

$$\mu_C=0,\mu_R=1-\mu_I \tag{50}$$

and

$$\mu_I(s) = \frac{\alpha q(1-\theta^*(s))}{\beta + \alpha q(1-\theta^*)}, \tag{51}$$

otherwise. Given the solutions (47) - (51), I find the repo spread in (61). $\square$

A.2 Extensions

Extension: Long-lived FIs

In this model, FIs whose investments mature, as well as FIs who go bankrupt, exit the economy and are replaced by an equal measure of new FIs. Here, I show the model's implications are robust to relaxing this assumption so that borrowers whose investment mature stay in the economy and choose to become either repo investors or cash investors until their investment opportunity arrives again.

To solve the model with long-lived FIs I need an assumption below to simplify the wealth distribution across the FIs. Even in a stationary equilibrium, the borrower's wealth at the investment maturity depends not only on how long the investment took but also on their wealth when the investment maturity arrived, making the problem intractable. So I assume FIs consume the proceeds except for their initial capital, making all the FIs start with the same wealth level.

Assumption 1. *FIs consume everything but their initial capital of 1 when their investment mature.*

I solve numerically both stationary equilibrium and transition path. The results are qualitatively the same and presented in 11.

The borrower's optimization problem now depends on the value function of repo investors/cash investors in the future. The levered and unlevered value function

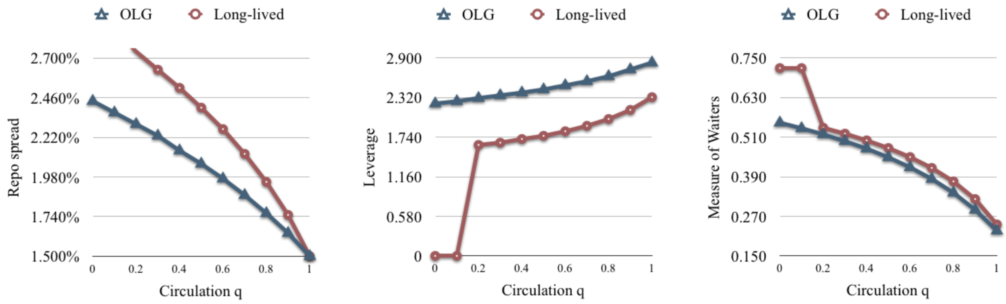


Figure 11: Comparing stationary equilibrium of the model (OLG) with long-lived FIs.

of the borrowers are therefore

$$V_{levered} \equiv \max_x E\left\{1 + x \cdot \hat{r}\left(\tau_\beta\right) + e^{-\rho\tau_\beta} \cdot \left(\max\left\{V_W, V_L\right\} - 1\right) \mid \tau_\beta \leq T_\delta\left(x\right)\right\} \quad (52)$$

$$V_{unlevered} \equiv E\left\{e^{-\rho\tau_\beta}\left(\frac{R}{\bar{p}} + \max\left\{V_W, V_L\right\} - 1\right)\right\} \quad (53)$$

Note that the higher value function of repo investor/cash investor (which is the same in equilibrium) increases the value function of unlevered value function. Also it tends to discourage borrowers from taking high leverage as it increases the losses they suffer upon bankruptcy.

A.3 Transition Path

Transition Path: Optimization Problems

Below I describe the borrower's optimization problem and the clearing of the secondary market along the transition path. The rest of model is not substantially changed from a stationary equilibrium by allowing dynamics. Here, I focus on the case where only the degree of circulation goes down from $q^- = 1$ to $q^+ < 1$, but no other parameter changes. The procedures are similar for other types of shocks.

Consider a borrower who receives an investment opportunity at $\forall t > 0$. He knows all the parameter values after the shock and faces possibly fluctuating haircuts and repo rates. As he chooses the size of investment once and cannot partially liquidate the project, he pays out of the buffer capital or is paid extra cash as the haircut fluctuate. Initially, he pledges optimal haircut capital x_t^* and borrows $\left(\frac{1}{h(t)}-1\right)x_t^*$. Suppose the haircut increases from $h(t)$ to $h(t+\Delta t)>h(t)$. To maintain the principal, the borrower needs to pledge $\Delta x = x(t+\Delta)-x(t)$ where

$$x(t+\Delta t) = \left[\frac{\frac{1}{h(t)}-1}{\frac{1}{h(t+\Delta t)}-1}\right]x_t^* \tag{54}$$

Denote by $\Delta(t,T)$ the cumulative discounted cash outflows due to changes in haircuts between t and $T>0$. Then we have

$$\Delta(t,T) = \Sigma_{\tau=t}^{T}\left\{e^{-\rho(\tau-t)}\cdot\left(\frac{\frac{1}{h(\tau)}-\frac{1}{h(\tau+\Delta t)}}{\frac{1}{h(\tau+\Delta t)}-1}\right)\right\} \tag{55}$$

If $h(t)$ is differentiable, $\Delta(t,T) = \int_t^T e^{-\rho(\tau-t)}\frac{\dot{h}(\tau)}{h(\tau)(1-h(\tau))}d\tau$.

Now, the borrower problem is similar to the one in stationary equilibrium except for the present value of net payoff per unit haircut capital $\hat{r}\left(\tau_\beta\right)$ and the solvency constraint T_δ. For a borrower with an investment opportunity at $\forall t>0$, we have

$$\hat{r}\left(\tau_\beta;t\right) \equiv e^{-\rho\tau_\beta}\left[1+\frac{1}{h(t)}\left(\frac{R}{\bar{p}}-1\right)\right]-\left\{\int_0^{\tau_\beta}e^{-\rho\tau}\left[\frac{1}{h(t)}-1\right]s(t+\tau)\,d\tau+\Delta\left(t,\tau_\beta\right)\right\}-1 \tag{56}$$

with the solvency constraint

$$\int_0^{T_\delta(x;t)}e^{-\rho\tau}\left[\frac{1}{h(t)}-1\right]s(t+\tau)\,d\tau+\Delta\left(t,t+T_\delta(x;t)\right)=\frac{1-x}{x} \tag{57}$$

Consider a borrower who is continuing his investment that started before the shock at $-t \in \left(-T_{\delta}^{-}, 0\right)$. The solvency constraint is now

$$\int_{0}^{-t+T_{\delta}^{-}} e^{-\rho\tau}\left(\frac{1}{h_0}-1\right) s^{-} d\tau =$$

$$\int_{0}^{-t+T_{\delta}\left(x^{-};-t\right)} e^{-\rho\tau}\left(\frac{1}{h_0}-1\right) s(\tau)\, d\tau + \Delta\left(0, -t+T_{\delta}\left(x^{-};-t\right)\right) \tag{58}$$

Here, the l.h.s. is the remaining buffer capital at the time the shock arrives. The r.h.s. is the total cash outflow due to the fluctuating repo rate and the haircut. The borrower initially chooses the maximum duration of T_{δ}^{-} of a stationary equilibrium, but the duration $T_{\delta}\left(x^{-};-t\right)$ changes due to the shock. Clearly, when the repo rate and/or haircut increases, the duration $T_{\delta}\left(x^{-};-t\right)$ becomes shorter than T_{δ}^{-}, meaning existing borrowers are forced into earlier bankruptcy.

Transition Path: Clearing the Secondary Market

First, let me explain how to solve the evolution of relative measures. For $\forall t > 0$, the measure of borrowers $\mu_I(t)$ changes is given

$$d\mu_I(t) = \alpha\left(\mu_W(t) + q\mu_L(t)\right) dt - \beta\mu_I(t)\, dt$$

$$-\int_{-\infty}^{t} \theta^{*}(\tau)\, \alpha\left(\mu_W(\tau) + \left[q^{-}\cdot 1_{\tau<0} + q\cdot 1_{\tau>0}\right]\mu_L(\tau)\right)\cdot 1_{\left\{\tau+T_{\delta}^{*}(\tau)=t\right\}} d\tau \tag{59}$$

The first two parts are similar to stationary equilibrium except that now q^{+} after the shock is lower than one. Fewer repo investors can circulate collateral and become borrowers. The third part is the outflow of the borrowers due to bankruptcy at t. The increasing haircut may force many borrowers into bankruptcy, resulting in the negative jump. The measure of repo investors and cash investors are straightforward.

Then we can calculate the total demand for assets and the supply of the secondary market. There is an excess supply when the supply of the secondary market is higher than total demand for the assets, in which case the price p should adjust downward to clear the market. At $\forall t > 0$, there is an excess supply in the secondary market if and only if

$$\alpha\left(\mu_{Wt} + q^{+}\mu_{Lt}\right)\cdot\frac{x_t^*}{h_t}dt < \int_{-\infty}^{t} e^{-\beta T_\delta^*(\tau)}\alpha\left(\mu_{W\tau} + q_\tau\mu_{L\tau}\right)\cdot\frac{x_\tau^*}{h_\tau}\cdot 1_{\left\{\tau+T_\delta^*(\tau)=t\right\}}d\tau \quad (60)$$

where the l.h.s. is the total demand, the flow of new borrowers multiplied by their leverage, and the r.h.s. is the supply of the secondary market, which is the aggregate quantity of assets liquidated by repo investors whose borrowers went bankrupt.

A.4 Bargaining Problem between Borrowers and Lenders

Here I consider the bargaining problem between borrowers and lenders in a repo contract. Suppose $s_l \leq s_b$ so that repos create a surplus. How lenders and borrowers divide the surplus depends on their bargaining power. In a competitive equilibrium with no search and matching frictions, the surplus division is determined by the market clearing condition.

A quick thought experiment demonstrates this bargaining problem. Suppose the equilibrium repo spread is greater than s_l ($s^* > s_l$) so that all FIs without opportunities prefer to lend. Also suppose there is sufficient cash in the system to meet the demand of all investors, and some FIs without opportunities have to be left as waiters. Then waiters are willing to receive $s_l + \epsilon$ for an arbitrary small $\epsilon > 0$ to attract borrowers and become lenders. Thus, if there exists a waiter with the spread s_l, the equilibrium repo spread is s_l in which case all the surplus is extracted by borrowers and FIs without opportunities are indifferent between lending and waiting. Now, what if there is not enough cash in the system, so that some investors cannot borrow at s_l? Then the investors who cannot borrow are willing to pay more to attract lenders, which drives up the repo spread. The repo spread will increase until all investors in the economy can borrow. However, the repo spread cannot be higher than the maximum spread s_b. If there are no waiters in the economy at the spread s_b, all the surplus is extracted by lenders and investors are indifferent between borrowing and not borrowing (if $s_b = \hat{s}$) or borrowing and lending (if $s_b = \hat{\hat{s}}$). To sum up, there are three cases, assuming that $s_l \leq s_b$:

1. $s^* = s_l$: The surplus is extracted by borrowers. All investors borrow and FIs without opportunities are indifferent between lending and waiting.

2. $s^* = s_b$: The surplus is extracted by lenders. Some investors borrow while other investors become lenders or make unlevered investments, whichever they prefer. All FIs without opportunities lend.

3. $s^* \in (s_l, s_b)$: The surplus is divided between borrowers and lenders so that the repo market clears and there are no waiters.

The division of surplus is determined by the equilibrium mass of cash investors, i.e., whether there is enough cash in the system to fund investors. It depends on the leverage that investors take, which in turn depends on the repo spread determined by the bargaining problem. Then, we can easily solve the bargaining problem between borrowers and lenders by examining the mass of waiters at different levels of repo spread $\mu_C(\cdot)$ as determined in (47) - (51).

Finally, consider the case when $s_l > s_b$. That is, the minimum spread for lending s_l is higher than the maximum spread for borrowing s_b. In this case, the repo contract does not create surplus, thus the repo markets shut down in equilibrium. The shadow price for the repo spread exists and equals s_l at which no investors borrow. So, we have one more type of equilibrium, in addition to the three considered above:

4. $s^* = s_l$: The repo markets shut down, as repos do not create surplus.

Lemma 3. *Then the equilibrium repo spread is*

$$s^* = \begin{cases} s_l & \text{if } \{s_l \leq s_b \text{ and } \mu_C(s_l) > 0\} \text{ or } s_l > s_b \\ s_b & \text{if } s_l \leq s_b \text{ and } \mu_C(s_b) = 0 \\ \max\left\{\mu_C^{-1}(0)\right\} \in (s_l, s_b) & \text{otherwise} \end{cases} \tag{61}$$

Also note that there are four types of equilibria. The repo markets open in Type 1 - Type 3 equilibria, where the maximum spread for investors to borrow is at least as high as the minimum spread for FIs without opportunities to lend; thus the repos create a surplus. The three types of equilbria differ in the division of the surplus created from repo contracts. In a Type 1 equilibrium, there is a positive mass of waiters and all the surplus is extracted by borrowers, leaving FIs without opportunities indifferent between lending and waiting. On the other hand, in a Type 2 equilibrium, the surplus is extracted by lenders and only some investors

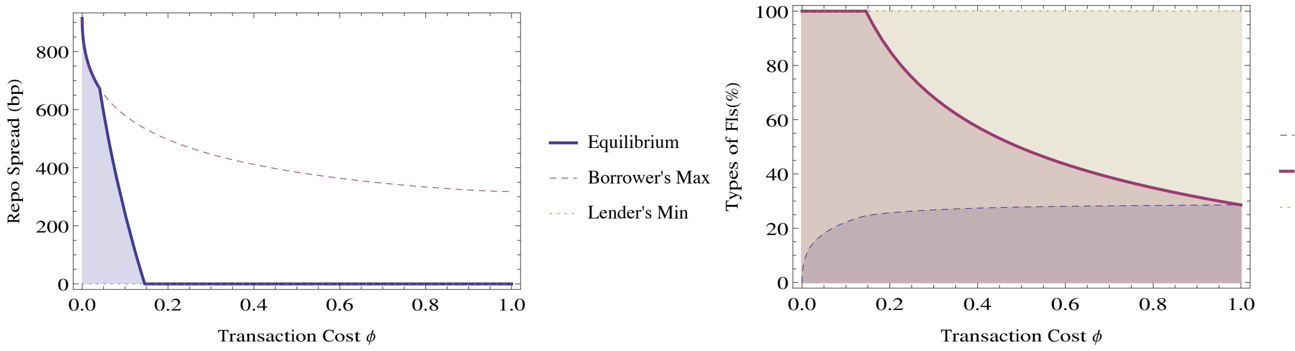


Figure 12: The figures plot the equilibrium repo spread and the masses of different types of FIs as the transaction cost parameter ϕ changes.

borrow from the repo market, leaving other investors to either not borrow or become lenders. In a Type 3 equilibrium, both borrowers and lenders are *more* better off by entering the repo contracts. Lastly, in a Type 4 equilibrium, there is no repo market as investors prefer not to borrow at the spread that lenders require.

I find that for an arbitrary set of parameters that describes the economy $\{q,\alpha,\beta,\sigma,\rho,R\}$, there is an interval for the transaction cost parameter ϕ in which a competitive stationary equilibrium is always of Type 1 (i.e., the repo market opens in an equilibrium and all investors can finance their leveraged investments). The following proposition summarizes this result.

Proposition 6. *For an arbitrary set of parameters $\{q,\alpha,\beta,\rho,R\}$, there always exists a lower bound of transaction cost $\check{\phi}(q,\alpha,\beta,\rho,R) \in [0,1]$, and an upper bound of the transaction cost $\hat{\phi}(q,\alpha,\beta,\rho,R) \in [0,1]$, such that a competitive stationary equilibrium with a set of parameters $\{q,\alpha,\beta,\rho,R,\phi\}$ is of Type 1 (i.e, $s_l \leq s_b$ and $\mu_W(s_l) > 0$) if and only if the transaction cost parameter ϕ satisfies*

$$\phi \in \left(\check{\phi}(q,\alpha,\beta,\rho,R), \hat{\phi}(q,\alpha,\beta,\rho,R)\right). \tag{62}$$

Moreover, with perfect circulation ($q = 1$), we have $\hat{\phi}(q, \alpha, \beta, \rho, R) = 1$ for $\forall \alpha, \beta, \rho$ and R. That is, the repo market always opens in equilibrium.

The intuition is as follows. In a stationary equilibrium the transaction cost equals the equilibrium haircut (I.26), as the demand for assets is always greater than the supply in the secondary market. Thus the interval for the transaction cost directly translates into the interval for the equilibrium haircut. As the transaction cost decreases ($h \searrow 0$), an individual borrower can take higher leverage while pledging the same amount of capital. The high leverage increases the demand for credit, saturating all the liquidity (cash) in the system. Thus with a very low level of the transaction cost, the system cannot provide enough cash and the borrowers have to compete for credit, resulting in either a Type 2 equilibrium (where lenders extract all the surplus) or a Type 3 equilibrium (where borrowers and lenders share the surplus). On the other hand, as the transaction cost increases, the amount of leverage an individual can take is very limited. With very low leverage ($h \nearrow 1$), borrowing may not be attractive enough to justify setting aside the buffer capital and paying the repo interest. An investor may prefer financing the investment with their cash only, resulting in a Type 4 equilibrium where the repo market completely shuts down.

Therefore, it is with the intermediate value of the transaction cost *and thus* the haircut) that a competitive stationary equilibrium is of Type 1. This implies that the type of market where collateral circulates among a limited number of FIs is suitable for collateral with intermediate liquidity. Collateral with very high liquidity may be financed by a larger pool of investors outside the system, while collateral with very low liquidity can only be financed by equity.

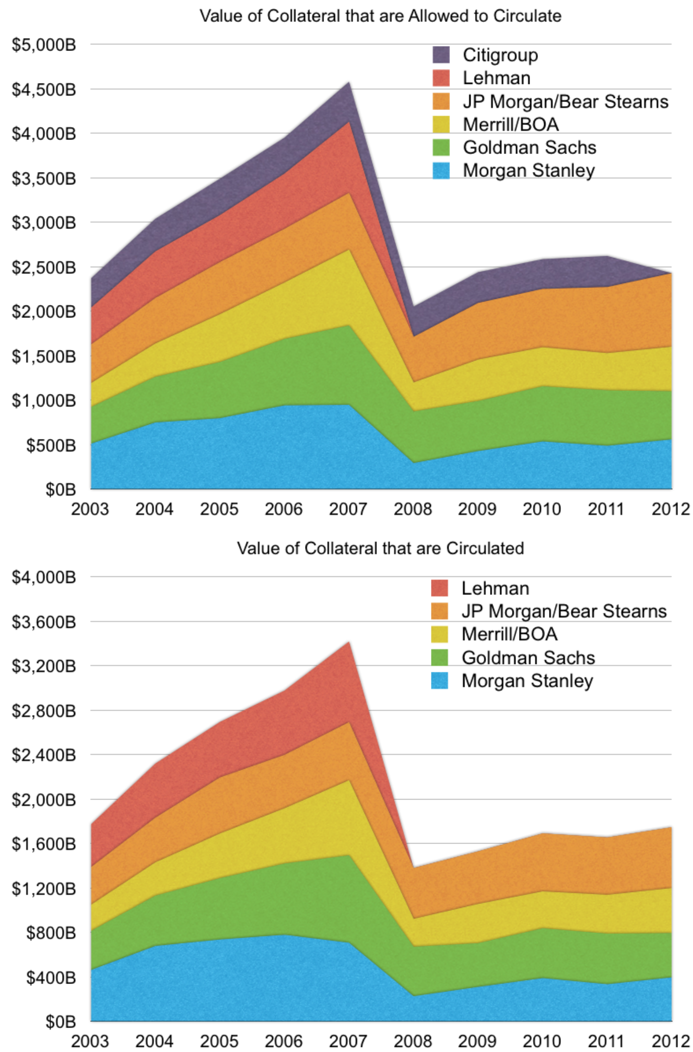


Figure 13: The data is from 10-K reports of major US broker-dealers: Morgan Stanley, Merrill Lynch/Bank of America, JP Morgan/Bear Stearns, Lehman Brothers, and Citigroup upon availability.

Chapter II: Heterogeneous Portfolio Constraints and Wealth Distribution Dynamics

1 Introduction

Empirical studies on household finance document that (i) many households in the United States do not participate in the stock market, (ii) the portfolio shares conditional on participation substantially differ across households and (iii) the portfolio shares exhibit inertia at the household level.[1] These facts raise two main questions. First, why do households behave as they do? Second, what will be the equilibrium implications of such behaviors for the aggregate economy?

In this paper, I attempt to answer the second question in a simple setup that reflects properties (i)-(iii). I extend the model of Basak and Cuoco (1998) to many types of agents including bankers and heterogeneous consumers. While bankers' portfolios are unconstrained, consumers' portfolios are constrained so that the shares of risky investment relative to wealth (portfolio shares[2]) have exogenous and heterogeneous upper bounds. In equilibrium, asset prices and the wealth distribution are jointly determined.

Generally, the wealth distribution should affect equilibrium asset prices. Gor-

[1] See Mankiw and Zeldes (1991), Vissing-Jorgensen (2003) and Brunnermeier and Nagel (2008) among others.

[2] This terminology follows Wachter and Yogo (2010)

man aggregation (the existence of a representative agent whose preferences are independent of the wealth distribution) only holds under special circumstances. For example, agents with CRRA preferences should have homogeneous relative risk aversion for Gorman aggregation to hold. Any departure from this condition implies significant dependence of asset prices on the wealth distribution.

Moreover, models with portfolio constraints exhibit equilibrium bubbles in both the stock and the bond prices. Hugonnier (2011) shows that there are equilibrium bubbles in both the stock and the bond in a model with two types of agents. Such properties carry over in my model with many types of agents. In particular, I show the existence of the stock bubble is equivalent to the existence of a consumer whose wealth relative to the bankers is expected to decline over time, which implies a diverging wealth distribution. Therefore in this model, not only does the wealth distribution affect asset prices, but also asset prices affect the wealth distribution as well.

My model exhibits an equilibrium Sharpe ratio (the equity premium per unit risk) that skyrockets when there is a negative aggregate shock to the economy. In fact, this countercyclicality of the Sharpe ratio is a quite general feature, driven by interaction of heterogeneity and the wealth effect.[3] Since agents who bear more risk (bankers in the model) will be hit harder by a negative shock, the market effectively becomes more risk averse and the Sharpe ratio goes up accordingly. While the Sharpe ratio depends on the entire distribution of wealth, I find one state variable that summarizes the wealth distribution regardless of the number of types of agents.

[3] Models with heterogeneous relative risk aversion as well as models with (heterogeneous) portfolio constraints exhibit countercyclical Sharpe ratios.

Furthermore, I ask when the Sharpe ratio would be more or less sensitive to aggregate shocks. I find that the Sharpe ratio is more sensitive in a bad economy compared to a good economy, (i.e., the countercyclicality of the Sharpe ratio is also countercyclical). Therefore the effect of serial bad shocks is greater than the sum of the effects of each shock. The frequency of shocks as well as their magnitudes matter for both the asset prices and wealth distribution.

In this model, the high Sharpe ratio in a bad economy is driven by the low real interest rate. When the Sharpe ratio is extremely high, the real interest rate goes negative, which implies high inflation. The government can help stabilize the economy by either transferring the wealth of consumers to bankers or relaxing the portfolio constraints of the consumers. It is more effective to act quickly rather than wait because the economy can become more vulnerable.

Transfering the wealth of consumers to bankers increases aggregate risk taking because bankers can take more risk than consumers whose portfolios are constrained. The Sharpe ratio decreases and the real interest rate increases immediately. It will be more effective to transfer the wealth of the most constrained consumers to bankers. This policy amplifies the trends of diverging wealth distribution.

Alternatively, the government can directly tackle consumers' portfolio constraints. Relaxing the constraints stabilizes not only the asset prices but also the wealth distribution. In the short run, the government can borrow from consumers and invest in the stock. When the economy stabilizes, the government can return the proceeds to the consumers. That way, consumers effectively invest more in the stock than their portfolio constraints allow them to do, when the Sharpe ratio is very high. In the long run, the government can mitigate the sources of the portfolio constraints.

This paper is silent about why consumers behave as they do and why their portfolios may be constrained. Basak and Cuoco (1998) point out some fixed information costs as one reason for stock market nonparticipation. ? show that nonparticipation may be a rational decision by ambiguity averse investors. The lack of insurance markets for idiosyncratic shocks may also force consumers to invest in the riskless bond. If the portfolio constraints are driven by such frictions, the government can relax the portfolio constraints by reducing them. For example, by making the financial system easier to access and/or providing social safety nets against idiosyncratic shocks. Instead, the portfolio constraints may simply reflect the preferences of consumers, in which case there is no role for the government.

1.1 Related Literature

This paper is related to several strands of literature. First of all, my model is based on the models of stock market nonparticipation. Basak and Cuoco (1998), motivated by Mankiw and Zeldes (1991) who find that only one-fourth US households own the stock, develop a general equilibrium model with stockholders and nonstockholders. They show that aggregate relative risk aversion of only 1.3 is enough to match the historical equity premium.[4] Chabakauri (2009) and Prieto (2010) extend their model to accommodate various types of portfolio constraints. Hugonnier (2011) show that there is an asset pricing bubble in their model, and therefore the relative wealth of nonstockholders decreases over time. My model is a direct extension of his model to many types of agents and to my best knowledge

[4]Their model has been subject to criticism because of the scanty wealth of nonstockholders. Nonstockholders own only around 10% of aggregate wealth of the US. (See Guvenen (2009)) This problem can be resolved in my model with heterogeneous portfolio constraints.

the first in this literature to have more than two types of agents.

Recent studies on household finance document more than stock market nonparticipation. Vissing-Jorgensen (2003) show that portfolio shares are heterogeneous conditional on stock market participation and the cross-sectional standard deviation of portfolio shares conditional on participation is around 30%. Brunnermeier and Nagel (2008) show that portfolio shares exhibit inertia for a given household. Their result is consistent with the exogenous portfolio share that is the equilibrium behavior of my model. Wachter and Yogo (2010) develop a model to explain why the portfolio shares are postively correlated with the wealth of households.

The dynamic relationship between borrowing constraints and the wealth distribution has been previously studied. Piketty (1997) studies the effect of credit rationing on the wealth distribution and finds that higher interest rates can be self-reinforcing through higher credit rationing and lower capital accumulation. Pawasutipaisit and Townsend (2011), using detailed data from an emerging market economy, study factors in achieving upward mobility in the wealth distribution. They find that the return on assets that depends on several demographic characteristics, such as the level of education. Their result indicates the significance of the portfolio constraints that households face. Piketty and Saez (2003), Vissing-Jorgensen (2009) and Parker and Vissing-Jorgensen (2010) study income and wealth inequality in the US. My model shows how heterogeneous portfolio constraints can contribute to the trends in wealth inequality.

Wang (1996) and Chan and Kogan (2002) study models with heterogeneous relative risk aversion. Wang (1996) models agents with CRRA preferences and show that less risk averse agents dominate the entire economy in the long run. Chan and Kogan (2002) model agents with "catching up with the Jones" preferences that equalize marginal utilities of agents with heterogeneous relative risk aversion. The wealth distribution in their model remains stationary. It turns out my model with heterogeneous portfolio constraints has the wealth distribution dynamics in-between the two.

Related to the recent financial crisis, He and Krishnamurthy (2008) extend the model of Basak and Cuoco (1998) to the case where the intermediary can directly invest in the stock and households can either invest in the bond or invest in the intermediary. They consider various government polices and find infusing equity capital into intermediaries is the most effective.

The paper proceeds as follows. The next section describes the model and defines an equilibrium. Section 3 solves for the equilibrium wealth dynamics. Section 4 discusses the model's implications for financial crises and government policies. Section 5 concludes.

2 The Model

The setup follows Basak and Cuoco (1998). I study a pure exchange economy with one risky stock S and one locally riskless bond B. I assume the horizon is finite ($[0,T]$). I omit the time subscript when it is clear from context. The model consists of $N+1$ (types of) agents including one group of bankers and N groups of consumers. Constrained agents have heterogeneous portfolio constraints that set the maximum portfolio shares (the proportion of wealth allocated to risky investment).[5]

[5] In Appendix, I show the model with a continuum of types and consider portfolio constraints that set the minimum portfolio shares.

2.1 Setup

Assets There is an exogenously given dividend process δ, which follows a diffusion with constant drift μ_δ and volatility σ_δ. It is the only source of uncertainty in the economy. The initial value δ_0 is normalized to 1.

$$\frac{d\delta_t}{\delta_t} = \mu_\delta dt + \sigma_\delta dZ_t \tag{II.1}$$

There are two assets; a risky stock S and a locally riskless bond B. The stock S can be interpreted as an index fund of all the stocks in the economy. It pays out the continuous dividend stream δ. The supply of stock is normalized to one share. An equilibrium stock price can be expressed as a geometric Brownian motion with drift μ and volatility σ. The parameters of stock price are determined in equilibrium and can be time-varying.

$$\frac{dS_t + \delta_t dt}{S_t} = \mu_t dt + \sigma_t dZ_t \tag{II.2}$$

The bond B is in zero net supply. The instantaneous interest rate r_t is determined in equilibrium. Note that r is the real interest rate denominated by the consumption good and can be negative due to inflation. The value of a money market fund involved in the safe bond B follows

$$\frac{dB_t}{B_t} = r_t dt \tag{II.3}$$

The Sharpe ratio, denoted κ_t, is defined as

$$\kappa_t = \frac{\mu_t - r_t}{\sigma_t}. \tag{II.4}$$

Let M denote the stochastic discount factor. The initial value M_0 is normalized to 1. The dynamics of M are completely determined by r and κ as

$$\frac{dM_t}{M_t} = -r_t dt - \kappa_t dZ_t \tag{II.5}$$

such that both MB and MS are nonnegative local martingales.[6]

Agents There are one group of bankers and N groups of consumers. Bankers and consumers have logarithmic preferences and time discount parameter ρ. The only source of heterogeneity among the agents is given by the maximum portfolio share. While bankers are unconstrained, consumers have exogeneous and heterogeneous upper bounds on portfolio shares that range from 0 to 1.

Bankers Since bankers are identical, we can think of a representative banker with the aggregate wealth of all bankers and call him "the banker". The banker solves for the optimal consumption and portfolio strategies (c^*_{Bt}, π^*_{Bt}) to maximize his lifetime utility subject to the budget constraint, where π_B denotes the amount of the consumption good invested in the stock. Suppose the banker's initial wealth is given by his stock holding $b_0 \in (0,1)$, i.e., $W_{B0} = b_0 \cdot S_0$.

Consumers There are N groups of consumers. For group $n \in \{1, \cdots, N\}$, there is a representative consumer for the group, called "consumer n". Consumer n's portfolio share is bounded above by $1 - \epsilon_n$ where $\epsilon_n \in (0,1]$. Assume $\{\epsilon_n\}_{n=1}^N$ is

[6] Throughout the paper, local martingales include martingales. Local martingales which are not martingales are referred as strict local martingales.

monotonically increasing. Thus, consumer 1 is the least constrained and consumer N is the most constrained.

Each consumer $n \in \{1, \cdots, N\}$ solves the optimal consumption and portfolio strategies (c_n^*, π_n^*) to maximize her lifetime utility subject to the budget constraint and the portfolio constraint. That is,

$$\max_{\{c_n, \pi_n\}} U(c_n) = E\left\{\int_0^T e^{-\rho t} \log(c_{nt})\, dt\right\} \tag{II.6}$$

subject to the budget constraint

$$W_{nt} = W_{n0} + \int_0^t (r_s \phi_{ns} + \mu_s \pi_{ns} - c_{ns})\, ds + \int_0^t \sigma_s \pi_{ns} dZ_s, \tag{II.7}$$

and the portfolio constraint

$$\pi_{nt} \leq (1 - \epsilon_n) W_{nt}, \tag{II.8}$$

where W_n denotes her wealth, and π_n and ϕ_n denote the amount invested in the stock and the bond respectively. Note that the portfolio constraints allow for short-sales. Suppose her initial wealth, W_{n0} is given by the stock holding $b_n \in (0,1)$, i.e. $W_{n0} = b_n \cdot S_0$.

Since the bond is in zero net supply, aggregate wealth is always the same as the stock price S, which implies,

$$b_0 + \Sigma_{n=1}^N b_n = 1. \tag{II.9}$$

Note that the initial portfolio share is equal to 1 for every agent so that portfolio constraints are violated for consumers at $t=0$. When the economy starts, consumers immediately adjust their portfolios to satisfy portfolio constraints.

2.2 Characterization of Equilibrium

Definition A competitive equilibrium of the economy is a price process $\{(B_t, S_t)\}_{t\in[0,T]}$ and a set of consumption and portfolio strategies for the banker and consumers $\left\{(c^*_{Bt}, \pi^*_{Bt}), (c^*_{nt}, \pi^*_{nt})_{n\in\{1,\cdots,N\}}\right\}_{t\in[0,T]}$ such that

1. Given a price process (B,S), the banker and consumers solve their unconstrained and constrained optimization problem described in 2.1 and the solutions are (c^*_B, π^*_B) and (c^*_n, π^*_n) for all $n = 1, \cdots, N$.

2. The consumption good market, stock market and bond market all clear.

Individual optimality The solution to the unconstrained optimization problem for logarithmic preference is widely known. The banker always consumes a fraction of his wealth and the fraction depends solely on his time preference. His portfolio share is determined by the Sharpe ratio and the volatility of the stock. For a constant $\gamma_B > 0$ and $\eta_t = \int_t^T e^{-\rho(s-t)} ds$,

$$c^*_{Bt} = \gamma_B \frac{e^{-\rho t}}{M_t} = \frac{W_B}{\eta_t} \text{ and } \pi^*_{Bt} = \frac{\kappa_t}{\sigma_t} W_{Bt} \tag{II.10}$$

Consumers 1 to N face an incomplete market because of the portfolio constraints. In general, it is not straightforward to solve for optimal strategies in an incomplete market. However, with convex portfolio constraints, as is the case in this paper, Cvitanić and Karatzas (1992) show that there exists a unique solution that corresponds to the solution of an unconstrained optimization problem in a slightly adjusted market.

The problem of solving for the strategies is transformed to that of finding the "market". Log investors' myopicity allows us to solve the transformed problem in a pointwise manner. Assuming that the adjusted market for each agent is known, I solve the constrained optimization problems. I denote the effective Sharpe ratio of the adjusted market for consumer n by κ_n. The effective stochastic discount factor for consumer n, denoted M_n, is then determined by κ_n and the bond price B as

$$M_{nt} = \frac{1}{B_t} \exp\left\{ -\int_0^t \kappa_{ns} dZ_s - \frac{1}{2} \int_0^t \left[\kappa_s^2 - (\kappa_s - \kappa_{ns})^2 \right] ds \right\} \tag{II.11}$$

Each consumer $n \in \{1, \cdots, N\}$ now solves her optimization problem as if she is unconstrained and faces the effective Sharpe ratio κ_n. Therefore, for a constant $\gamma_n > 0$,

$$c_{nt}^* = \gamma_n \frac{e^{-\rho t}}{M_{nt}} = \frac{W_{nt}}{\eta_t} \text{ and } \pi_{nt}^* = \frac{\kappa_{nt}}{\sigma_t} W_{nt} \tag{II.12}$$

Market clearing conditions I substitute individual optimal solutions (II.10) and (II.12) into the market clearing conditions for the consumption good, the stock and the bond. First, the consumption goods market clears.

$$c_B^* + \Sigma_{n=1}^N c_n^* = \frac{W_B}{\eta} + \Sigma_{n=1}^N \frac{W_n}{\eta} = \frac{\overline{W}}{\eta} = \delta \tag{II.13}$$

where $\overline{W} = W_0 + \Sigma_{n=1}^N W_n$.

The stock and bond markets clear.

$$\pi_B^* + \Sigma_{n=1}^N \pi_n^* = \frac{\kappa}{\sigma} W_B + \Sigma_{n=1}^N \frac{\kappa_n}{\sigma} W_n = S \tag{II.14}$$

$$\phi_B^* + \Sigma_{n=1}^N \phi_n^* = 0 \tag{II.15}$$

Since the bond is in zero net supply and from (II.13), we have

$$S = \bar{W} = \eta\delta \tag{II.16}$$

By applying Ito's lemma to the stock price $S = \eta\delta$, we get

$$\mu_t \equiv \rho + \mu_\delta \text{ and } \sigma_t \equiv \sigma_\delta \tag{II.17}$$

Moreover, from (II.14) and (II.16), we have

$$\kappa \frac{W_B}{\bar{W}} + \Sigma_{n=1}^N \kappa_n \frac{W_n}{\bar{W}} = \sigma_\delta \tag{II.18}$$

which indicates that the wealth-weighted average of the effective Sharpe ratios is independent of the portfolio constraints and equal to the volatility of the dividend, σ_δ.

The equilibrium Sharpe ratio Consider the optimal portfolio strategy π_n^* for consumer n. Since her optimal portfolio satisfies the portfolio constraints, from (II.8) and (II.12), we have $\kappa_n \leq (1-\epsilon_n)\sigma_\delta$. Since $\epsilon_n > 0$ for $\forall n$, effective Sharpe ratios are smaller than σ_δ for every consumer. Since the wealth weighted average of the Sharpe ratios is always σ_δ and the effective Sharpe ratio for every consumer is smaller than σ_δ, the market Sharpe ratio should be strictly greater than σ_δ. Therefore the portfolio constraint always binds for every consumer $n \in \{1, \cdots, N\}$, i.e., for $\forall n$,

$$\kappa_n = (1-\epsilon_n)\sigma_\delta. \tag{II.19}$$

By substituting κ_n's into (II.18) and from (II.17), we get

$$\kappa = \left(1 + \Sigma_{n=1}^{N} \epsilon_n \frac{W_n}{W_B}\right) \sigma_\delta, \tag{II.20}$$

$$r = \rho + \mu_\delta - \left(1 + \Sigma_{n=1}^{N} \epsilon_n \frac{W_n}{W_B}\right) \sigma_\delta^2, \tag{II.21}$$

Note that the wealth ratios between the banker and consumers are crucial in equilibrium. The effects of portfolio constraints are amplified when the banker is relatively poor compared to consumers. In the next section, I solve for the wealth dynamics of the banker and consumers in the economy and fully describe an equilibrium.

3 Dynamics of Wealth Distribution

Denote by λ_n the relative wealth of consumer n compared to the banker. The relative wealth is well-defined as long as the banker has positive wealth, which is the case because of the Inada conditions. From the optimal consumption strategies of the banker and consumers (II.10) and (II.12), we have

$$\lambda_n = \frac{W_n}{W_B} = \frac{c_n^*}{c_0^*} = \gamma_n \frac{M}{M_n}, \tag{II.22}$$

where a positive constant γ_n is recycled to denote γ_n / γ_B. From (II.5) and (II.11), we have for $\forall n \in \{1, \cdots, N\}$,

$$\frac{M}{M_n} = \exp\left\{-\int_0^t (\kappa_s - \kappa_{ns})\, dZ_s - \frac{1}{2}\int_0^t (\kappa_s - \kappa_{ns})^2\, ds\right\} \tag{II.23}$$

and from (II.22), we have

$$d\left(\frac{M}{M_n}\right) = -\left(\kappa - \kappa_n\right)\frac{M}{M_n}dZ \text{ and } d\lambda_n = -\left(\kappa - \kappa_n\right)\lambda_n dZ, \tag{II.24}$$

which implies λ_n is a nonnegative local martingale, thus a supermartingale, i.e. $E_t\{\lambda_s\} \leq \lambda_t$ for $s \geq t$.

From (II.20) and the effective Sharpe ratios for the consumers, we get

$$\kappa - \kappa_n = \left(\epsilon_n + \Sigma_{l=1}^{N}\frac{\epsilon_l W_l}{W_B}\right)\sigma_\delta = \left[\epsilon_n\left(1+\lambda_n\right) + \Sigma_{l\neq n}\epsilon_l\lambda_l\right]\sigma_\delta. \tag{II.25}$$

Substitute (II.25) into (II.24) and rearrange to get

$$d\lambda_n = \left[\left(1+\lambda_n\right)\left(-\epsilon_n\lambda_n\right) + \Sigma_{l\neq n}\lambda_n\left(-\epsilon_l\lambda_l\right)\right]\sigma_\delta dZ. \tag{II.26}$$

The dynamics of wealth ratio λ_n of consumer n to the banker depends on not only λ_n, but also the wealth ratios of other consumers to the banker as shown in the matrix

$$\begin{bmatrix} d\lambda_1 \\ d\lambda_2 \\ \vdots \\ d\lambda_N \end{bmatrix} = \begin{bmatrix} 1+\lambda_1 & \lambda_1 & \cdots & \lambda_1 \\ \lambda_2 & 1+\lambda_2 & \cdots & \lambda_2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_N & \lambda_N & \cdots & 1+\lambda_N \end{bmatrix}\begin{bmatrix} -\epsilon_1\lambda_1 \\ -\epsilon_2\lambda_2 \\ \vdots \\ -\epsilon_N\lambda_N \end{bmatrix}\sigma_\delta dZ. \tag{II.27}$$

The initial value is given by the initial allocation of wealth and for $\forall n$,

$$\lambda_n\left(0\right) = \frac{W_{n0}}{W_{B0}} = \frac{b_n}{b_0}. \tag{II.28}$$

Proposition 1 The equilibrium wealth distribution is given by the realtive wealth processes $\{\lambda_n(t)\}_{n=1}^N$. The relative wealth processes evolve according to (II.27) with the initial value (II.28).

Whether the relative wealth process is a martingale is crucial for the dynamics of wealth distributions. It is also the key in the existence of the stock bubble. In the Appendix, I show that the existence of a stock price bubble is equivalent to the existence of a consumer whose relative wealth compared to the banker decreases over time in expectation. I also show that the relative wealth processes are not martingales, but strict local martingales. Therefore the stock price has a bubble and the consumers are expected to become poorer compared to the banker over time.

Lemma 1 There exists a stock price bubble if and only if there exists a relative wealth process that is not a martingale but a strict local martingale.

Proof. See Appendix. □

Lemma 2 The relative wealth process λ_n for $\forall n$ are not martingales but strict local martingales.

Proof. See Appendix. □

Figure 1 shows the direct effects of portfolio constraints on the wealth distribution. Lorenz curves plot how much of aggregate wealth is held by a given proportion of population. A linear Lorenz curve implies perfectly equal wealth distribution. I assume the banker and consumer have a perfectly equal wealth distribution and plot the trends in Lorenz curves for each distribution of portfolio constraints. For every case, Lorenz curves become more convex and the wealth inequality

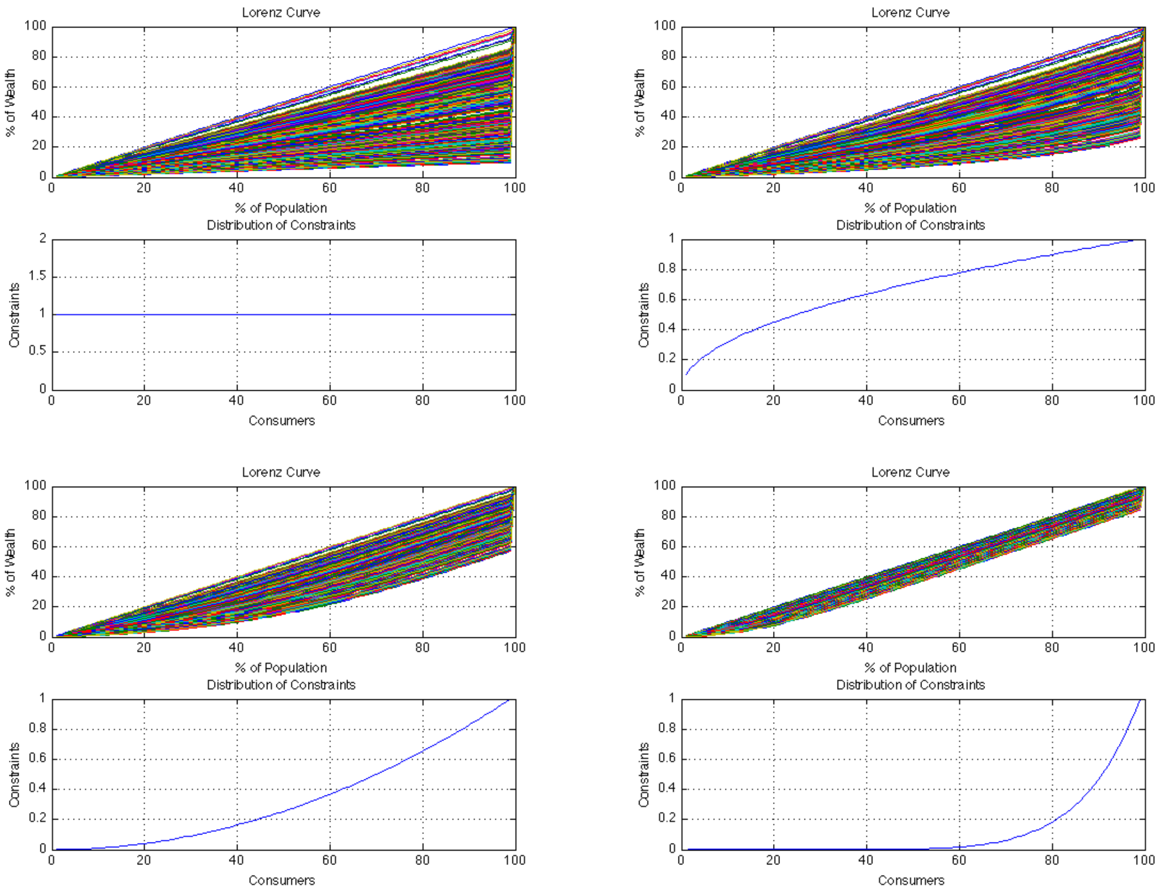


Figure II.1: Lorenz curves with the distribution of portfolio constraints
The top left panel is when almost no consumer participates in the stock market. The bottom right panel is when more than 50 percent of consumers have almost no portfolio constraints.

increases over time. However, the extent to which the wealth distribution diverges substantially depends on the portfolio constraints. As the portfolio constraints for consumers are tighter, wealth inequality increases much more quickly.

While equilibrium prices depend on all values of λ_n's, it is sufficient to know the constraint-weighted sum of λ_n's. Define Λ_k for $\forall k = 0, 1, \cdots$ as

$$\Lambda_k \equiv \Sigma_{n=1}^{N} \epsilon_n^k \lambda_n. \tag{II.29}$$

Note that given $\{\lambda_n\}_{n=1}^N$, $\{\Lambda_k\}_{k=0}^\infty$ is monotoncially decreasing because $\epsilon_n \in (0,1]$. Since ϵ_i^k approaches zero for any given $\epsilon_i < 1$, I obtain $\lim_{k\to\infty} \Lambda_k = \lambda_N$. The law of motion for Λ_k for $\forall k$ is determined recursively as

$$d\Lambda_k = -(\Lambda_{k+1} + \Lambda_k \Lambda_1)\sigma_\delta dZ. \tag{II.30}$$

This analysis extends simply to the case of a continuum of types of consumers.

Proposition 2. The equilibrium Sharpe ratio and real interest rate are determined as below.

$$\kappa = (1 + \Lambda_1)\sigma_\delta, \tag{II.31}$$

$$r = \rho + \mu_\delta - (1 + \Lambda_1)\sigma_\delta^2. \tag{II.32}$$

The initial value and law of motion for Λ_1 are given by (II.30). Equilibrium wealth dynamics for every agent $n \in \{1, \cdots, N\}$ are determined by (II.27) with given initial conditions (II.28).

Effective risk aversion As in Prieto (2010), consumers' constrained optimal portfolios can be interpreted as a result of their heterogeneous relative risk aversion. Note that the optimal portfolio of consumer n is the same as that of a CRRA investor with relative risk aversion θ_{nt} at a given time t, where

$$\theta_{nt} = \frac{\kappa_t}{\kappa_n} = \frac{1 + \Lambda_{1t}}{1 - \epsilon_n}. \tag{II.33}$$

The relative risk aversion, however, changes over time as the Sharpe ratio changes over time. Consumers behave as if they are more risk averse when the economy is bad. This allows me to compare my model with models with heterogeneous relative risk aversion. Wang (1996) model agents with CRRA preferences and show that less risk averse agents deterministically dominate the entire economy in the long run. Chan and Kogan (2002) model agents with "catching up with the Jones" preferences that equalize marginal utilities of agents with heterogeneous relative risk aversion. The wealth distribution in their model remains stationary. However, in my model, the bankers who are unconstrained and effectively the least risk averse are expected to become wealthier compared to the consumers, but only in expectation. The trends in wealth distribution in my model can be said to be exactly in between of those two models.

By summing up the effective risk tolerance (inverse of absolute risk aversion), I can get the effective market risk aversion

$$\theta = \frac{\bar{W}}{W_B + \Sigma_{n=1}^{N} \frac{W_n}{\theta_n}} = 1 + \Lambda_1 = \frac{\kappa}{\sigma_\delta}. \tag{II.34}$$

The market risk aversion θ is always greater than 1 and countercyclical.

4 Crises and Government Bailouts

This model has the interesting feature that the equilibrium Sharpe ratio and equity premium skyrocket when the economy is very bad. In fact, it is a quite general feature driven by the interaction of heterogeneity and the wealth effect. When there is a negative shock, agents with higher risk capacity suffer more and become relatively poorer. The aggregate risk capacity of the market shrinks, thus the Sharpe ratio skyrockets. In concurrent research, He and Krishnamurthy (2008) study such dynamics during crises in a model where only financial intermediaries can invest in the stock market. This model also finds the countercyclicality of Sharpe ratio, i.e., high Sharpe ratios in recessions and low Sharpe ratios in booms. As shown in the previous section, the equilibrium Sharpe ratio depends on the state variable Λ_1 that summarizes the wealth distribution regardless of the number of agents. The state variable Λ_1 is the constraint-weighted sum of wealth ratios of consumers and the banker. To see the cyclical properties of the Sharpe ratio, consider

$$(d\kappa) \cdot (dZ) = -\left(\Lambda_2 + \Lambda_1^2\right) \sigma_\delta^2 dt < 0 \tag{II.35}$$

which indicates the countercyclicality of the Sharpe ratio. Due to logarithmic preferences of consumers and the banker, the stock price has a constant drift and volatility. Instead, the countercyclicality of the Sharpe ratio is driven by the interest effect. In the bad economy, i.e., the aggregate dividend realization is very low, the real interest rate drops steeply, indicating high inflation.

Then when would the economy be more or less sensitive to bad shocks? To answer this question, I explore the cyclical property of such countercyclicality. The model's mathematical tractability allows us to directly caculate

$$d\left(\left|\frac{(d\kappa) \cdot (dZ)}{dt}\right|\right) \cdot (dZ) = -\left[\Lambda_3 + \Lambda_1 \Lambda_2 + 2\left(\Lambda_2 + \Lambda_1^2\right) \Lambda_1\right] \sigma_\delta^3 dt < 0, \tag{II.36}$$

which indicates the countercyclicality of the countercyclicality. In other words, the economy is more vulnerable to bad shocks when the economy is already bad. Therefore the effect of serial bad shocks is bigger than the sum of effects of individual shocks. Not only the size of shocks but also the frequency of the shocks matter. To my best knowledge, this paper is first to theoretically show this effect.

Corollary 1. The equilibrium Sharpe ratio κ is countercyclical and the extent to which κ is countercyclical is also countercyclical.

Now I consider potential government policies to stabilize the economy with a high Sharpe ratio and inflation. To reduce the Sharpe ratio and increase the real interest rate, the government should decrease Λ_1 because of (II.31). From

$$\Lambda_1 = \Sigma_{n=1}^{N} \epsilon_n \frac{W_n}{W_B}, \tag{II.37}$$

we can see the government needs to increase the banker's wealth ($W_B \uparrow$), decrease the wealth of consumers ($W_n \downarrow$), or relax the portfolio constraints for the consumers ($\epsilon_n \downarrow$). Moreover, it is more effective to act quickly rather than wait because the economy becomes more vulnerable to bad shocks (due to the countercyclical countercyclicality of the Sharpe ratio).

Transfering the wealth of consumers to bankers can decrease the consumers' wealth and increase the banker's wealth. This is to bail out bankers by taxing consumers. The transfer increases aggregate risk capacity and decreases the Sharpe ratio accordingly. Moreover, for a given size transfer, it is most effective to transfer the wealth of the most constrained consumers to bankers. Such a transfer of wealth from consumers to the banker amplifies the trends of diverging wealth distribution. Alternatively, the government can directly tackle consumers' portfolio constraints. Relaxing the constraints stabilizes not only the asset prices but also the wealth distribution. In the short run, the government can borrow from consumers and invest in the stock. Borrowing from consumers can be done by taxing them or issuing government debts. Consumers may view the government debts as equivalent to the banker's debt.[7] When the economy stabilizes, the government can return the proceeds to the consumers directly or indirectly by reducing taxes. That way, consumers effectively invest more in the stock than their portfolio constraints allow them to do, especially when the Sharpe ratio is very high.

In the long run, the government can mitigate the sources of the portfolio constraints. This requires a rigorous investigation on why consumers behave as they do, which is beyond the scope of this paper. Previous studies on stock market nonparticipation sugggest some fixed information costs (Basak and Cuoco (1998)) and ambiguity aversion (?). Or, the lack of insurance markets for idiosyncratic shocks may also force consumers to invest in the riskless bond (the precautionary saving motive). If the portfolio constraints are driven by such frictions, the government can relax the portfolio constraints by, for example, making the financial system more transparent and easier (reducing information costs), regulating unlikely events (reducing ambiguity), and/or providing social safety nets against idiosyncratic shocks (reducing precautionary saving motives). However, the portfolio constraints may simply reflect the preferences of consumers as heterogeneous effective relative risk aversion, in which case there is no role for the government.

[7] The probability of banker's default is zero in a physical probability measure, because the banker has log utility. Therefore both the bankers' and government debt here are considered riskless.

5 Conclusion

I study a general equilibrium model of stock market nonparticipation and heterogeneous portfilo constraints. In equilibrium, the asset prices and wealth distribution are jointly determined. I find that the Sharpe ratio is countercyclical and the countercyclicality is also countercyclical. I also show the relative wealth of consumers to bankers is expected to decrease over time. I discuss the implications of the model for financial crises and government bailouts.

It is left for future study to explain why consumers behave as they do. Whether consumers' portfolio constraints are driven by various frictions or simply reflect their preferences is crucial in the policy perspective. Moreover, this model is a pure exchange economy. It will be interesting to study the interaction between portfolio constraints and production in a macroeconomic model.

A Appendix

Proof for Lemma 1 Define the fundamental value F of the stock S as the discounted present value of the future cash flow. Then from (II.24), we have

$$F_t = E_t\left\{\int_t^T \frac{M_s}{M_t}\delta_s dt\right\} = \delta_t E_t\left\{\int_t^T e^{-\rho(s-t)}\left(\frac{1+\Sigma_{n=1}^N \lambda_{ns}}{1+\Sigma_{n=1}^N \lambda_{nt}}\right) ds\right\} \tag{38}$$

From (II.10), (II.16), the stock price is

$$S_t = \delta_t E_t\left\{\int_t^T e^{-\rho(s-t)} ds\right\} \tag{39}$$

and there is a stock bubble ($S_t > F_t$) if and only if

$$\int_t^T e^{-\rho(s-t)} \left(\frac{\Sigma_{n=1}^N \left(\lambda_{nt} - E\left\{\lambda_{ns}\right\}\right)}{1+\Sigma_{n=1}^N \lambda_{nt}} \right) ds > 0 \tag{40}$$

Therefore, there exists a stock price bubble if and only if there exists a relative wealth process that is a strict local martingale. $\square$

Proof for Lemma 2 I want to prove that λ_n is not a martingale by the Feller's explosion test that is used in Heston, Loewenstein, and Willard (2007) for all $n \in \{1, \cdots, N\}$, . The spirit of the proof is first to claim that the process is a martingale and the risk neutral measure is equivalent. Then I show that the risk neutral measure is not equivalent to (or absolutely continuous with respect to) the physical measure, therefore λ_n is not a martingale.

Proof. Consider the candidate risk neutral measure and denote its Brownian motion by dZ^Q. Then by Girsanov's theorem,

$$d\lambda_n = -\left(\epsilon_n + \Lambda_1\right) \lambda_n \sigma_\delta \left(-\kappa dt + dZ^Q\right) \tag{41}$$

$$= \left(\epsilon_n + \Lambda_1\right)\left(1+\Lambda_1\right) \sigma_\delta^2 \lambda_n dt - \left(\epsilon_n + \Lambda_1\right) \sigma_\delta \lambda_n dZ^Q \tag{42}$$

Denote the drift and volatility of λ_n by a and b as

$$a = \left(\epsilon_n + \Lambda_1\right)\left(1+\Lambda_1\right) \sigma_\delta^2 \lambda_n \tag{43}$$

$$b = -\left(\epsilon_n + \Lambda_1\right) \sigma_\delta \lambda_n \tag{44}$$

Then

$$\frac{a}{b^2} = \frac{1}{\lambda_n}\left(\frac{1+\Lambda_1}{\epsilon_n+\Lambda_1}\right) \tag{45}$$

where

$$1 \leq \frac{1+\Lambda_1}{\epsilon_n+\Lambda_1} \leq \frac{1}{\epsilon_n} \tag{46}$$

Consider the scale measure $p_c(x)$ defined by

$$p_c(x) = \int_c^x \exp\left[-2\int_c^\eta \frac{a(\phi)}{b(\phi)^2}d\phi\right]d\eta \tag{47}$$

Then $p_c'(x)$ is bounded as

$$\frac{e^{-2/\epsilon_n}}{c}x \leq p_c'(x) = \exp\left[-2\int_c^x \frac{a(\phi)}{b(\phi)^2}d\phi\right] \leq \frac{e^{-2}}{c}x \tag{48}$$

Therefore the speed measure $v_c(x)$ defined by

$$v_c(x) = \int_c^x p_c'(y)\int_c^y \frac{2}{p_c'(z)\,b(z)^2}dzdy \tag{49}$$

is bounded above as

$$k_1\int_c^x y\int_c^y \frac{1}{zb(z)^2}dzdy \leq v_c(x) \leq k_2\int_c^x y\int_c^y \frac{1}{zb(z)^2}dzdy \tag{50}$$

for some constant $k_1, k_2 > 0$. $v_c(x)$ does not explode as $x \to 0$ since $y\int_c^y \frac{1}{zb(z)^2}dz$ is unbounded as $y \to 0$. However, $v_c(x)$ explodes as $x \to \infty$ since $\int_c^x y\int_c^y \frac{1}{zb(z)^2}dzdy$ is bounded as $x \to \infty$. This implies that λ_n reaches infinity at a finite time with a positive probability under the risk neutral measure, while it doesn't under the physical measure because preferences satisfy Inada conditions and the dividend never hits zero. Therefore the two measures are not equivalent, thus λ_n is not a martingale for every $n \in \{1, \cdots, N\}$.$\square$ $\square$

A continuum of agents The model can be easily extended to the case with a continuum of constrained agents. I assume a mass of constrained agents normalized to one. Constrained agents are indexed by i for $i \in [0,1]$. Consider an arbitrary integrable function $\epsilon : [0,1] \to (0,1]$. Define a portfolio constraint for Agent i as $\pi_i \leq (1-\epsilon(i)) W_i$. Recycle the notations $\{\Lambda_k\}_{k=0}^{\infty}$ by defining Λ_k as

$$\Lambda_k = \int_0^1 \epsilon(i)^k \lambda_i di \tag{51}$$

I can solve for $\{d\Lambda_k\}_{k=0}^{\infty}$ by multiplying $\epsilon(i)^k$ and integrating over i. For $\forall k \geq 0$, I obtain the exactly same dynamics (II.30) and the Sharpe ratio is also the same as (II.31).

Minium portfolio share constraints Now I consider portfolio constraints which force agents to take more risks than they would otherwise. For simplicity, consider three types of agents: unconstrained Agent 0, constrained with maximum portfolio share Agent n and constrained with minimum portfolio share Agent m. For some $\epsilon_n, \epsilon_m > 0$, the portfolio share of Agent m is bounded above by $1-\epsilon_n$ and that of Agent n is bounded below by $1+\epsilon_m$. Denote the effective Sharpe ratios for Agent m and n by κ_m and κ_n respectively.

Solving for their optimal portfolio, I get $\kappa_m > \sigma_\delta$ and $\kappa_n < \sigma_\delta$ and their constraints may not be binding. Hence I need to consider three cases as below.

$$
\begin{cases}
\text{Case 1:} & \kappa < (1-\epsilon_n)\,\sigma_\delta \\
\\
\text{Case 2}: & (1-\epsilon_n)\,\sigma_\delta \leq \kappa < (1+\epsilon_m)\,\sigma_\delta \\
\\
\text{Case 3:} & \kappa \geq (1+\epsilon_m)\,\sigma_\delta
\end{cases} \tag{52}
$$

In Case 1, the portfolio constraint of Agent m binds and that of Agent n doesn't. Then I get $\kappa_m = (1+\epsilon_m)\,\sigma_\delta$ and $\kappa_n = \kappa$. The equilibrium Sharpe ratio is

$$
\kappa = \sigma_\delta \left(1 - \frac{\epsilon_m W_m}{W_0 + W_n}\right) \tag{53}
$$

Since $\kappa < (1-\epsilon_n)\,\sigma_\delta$, this implies

$$
\epsilon_m W_m > \epsilon_n W_n + \epsilon_n W_0 \tag{54}
$$

Equilibrium wealth dynamics are obtained as

$$
\begin{bmatrix} d\lambda_m \\ \\ d\lambda_n \end{bmatrix} = \begin{bmatrix} 1+\lambda_m & \lambda_m \\ \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \epsilon_m \lambda_m \\ \\ -\epsilon_n \lambda_n \end{bmatrix} \sigma_\delta dZ \tag{55}
$$

In Case 2, the portfolio constraints for both Agent m and n bind. The equilibrium Sharpe ratio is

$$
\kappa = \sigma_\delta \left(1 + \frac{\epsilon_n W_n}{W_0} - \frac{\epsilon_m W_m}{W_0}\right) \tag{56}
$$

where

$$
\epsilon_n W_n - \epsilon_m W_0 < \epsilon_m W_m \leq \epsilon_n W_n + \epsilon_n W_0 \tag{57}
$$

Equilibrium wealth dynamics are obtained as

$$\begin{bmatrix} d\lambda_m \\ d\lambda_n \end{bmatrix} = \begin{bmatrix} 1+\lambda_m & \lambda_m \\ \lambda_n & 1+\lambda_n \end{bmatrix} \begin{bmatrix} \epsilon_m \lambda_m \\ -\epsilon_n \lambda_n \end{bmatrix} \sigma_\delta dZ. \tag{58}$$

In Case 3, the portfolio constraint of Agent m does not bind and that of Agent n is binds. The equilibrium Sharpe ratio is

$$\kappa = \sigma_\delta \left(1 + \frac{\epsilon_n W_n}{W_0 + W_m}\right), \tag{59}$$

where

$$\epsilon_m W_m \leq \epsilon_n W_n - \epsilon_m W_0 \tag{60}$$

Equilibrium wealth dynamics are again obtained similarly.

$$\begin{bmatrix} d\lambda_m \\ d\lambda_n \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \lambda_n & 1+\lambda_n \end{bmatrix} \begin{bmatrix} \epsilon_m \lambda_m \\ -\epsilon_n \lambda_n \end{bmatrix} \sigma_\delta dZ. \tag{61}$$

Bibliography

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, 2013, Systemic risk and stability in financial networks, *Working paper*.

Acharya, Viral V., Douglas Gale, and Tanju Yorulmazer, 2011, Rollover risk and market freezes, *The Journal of Finance* 66, 1177–1209.

Acharya, Viral V., Denis Gromb, and Tanju Yorulmazer, 2012, Imperfect competition in the interbank market for liquidity as a rationale for central banking, *American Economic Journal: Macroeconomics* 4, 184–217.

Acharya, Viral V., and David Skeie, 2011, A model of liquidity hoarding and term premia in inter-bank markets, *Journal of Monetary Economics* 58, 436 – 447.

Acharya, Viral V., and Tanju Yorulmazer, 2007, Too many to fail—an analysis of time-inconsistency in bank closure policies, *Journal of Financial Intermediation* 16, 1 – 31.

Adrian, Tobias, Brian Begalle, Adam M. Copeland, and Antoine Martin, 2012, Repo and securities lending, .

Adrian, Tobias, and Hyun Song Shin, 2010, Liquidity and leverage, *Journal of Financial Intermediation* 19, 418 – 437 Risk Transfer Mechanisms and Financial Stability.

——— , 2011, Procyclical leverage and value-at-risk, *FRB of New York Staff Report*.

Allen, Franklin, and Douglas Gale, 1994, Limited market participation and volatility of asset prices, *The American Economic Review* 84, 933–955.

——— , 2000, Financial contagion, *Journal of Political Economy* 108, 1–33.

——— , 2005, From cash-in-the-market pricing to financial fragility, *Journal of the European Economic Association* 3, 535–546.

Aragon, George O., and Philip E. Strahan, 2012, Hedge funds as liquidity providers: Evidence from the lehman bankruptcy, *Journal of Financial Economics* 103, 570 – 587.

Azariadis, Costas, James Bullard, and Bruce D. Smith, 2001, Private and public circulating liabilities, *Journal of Economic Theory* 99, 59 – 116.

Bagehot, Walter, 1878, *Lombard Street: a description of the money market* (C. Kegan Paul).

Basak, S, and D Cuoco, 1998, An equilibrium model with restricted stock market participation, *Review of Financial Studies* 11, 309–341.

Bhattacharya, S., and Douglas Gale, 1987, Preference shocks, liquidity, and central bank policy, in *New Approaches to Monetary Economics* (Cambridge University Press).

Bottazzi, Jean-Marc, Jaime Luque, and Mário R. Páscoa, 2012, Securities market theory: Possession, repo and rehypothecation, *Journal of Economic Theory* 147, 477 – 500 <ce:title>Issue in honor of David Cass</ce:title>.

Brunnermeier, Markus K., and Stefan Nagel, 2008, Do wealth fluctuations generate time-varying risk aversion? micro-evidence on individuals, *American Economic Review* 98, 713–36.

Brunnermeier, Markus K., and Lasse Heje Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

Caballero, Ricardo J., and Arvind Krishnamurthy, 2008, Collective risk management in a flight to quality episode, *The Journal of Finance* 63, 2195–2230.

Caballero, Ricardo J., and Alp Simsek, 2009, Fire sales in a model of complexity, .

Carr, Peter, 1998, Randomization and the american put, *Review of Financial Studies* 11, 597–626.

Chabakauri, Georgy, 2009, Asset pricing in general equilibrium with constraints, .

Chan, Yeung Lewis, and Leonid Kogan, 2002, Catching up with the joneses: Heterogeneous preferences and the dynamics of asset prices, *Journal of Political Economy* 110, 1255–1285.

Cifuentes, Rodrigo, Gianluigi Ferrucci, and Hyun Song Shin, 2005, Liquidity risk and contagion, *Journal of the European Economic Association* 3, 556–566.

Copeland, Adam M., Antoine Martin, and Michael Walker, 2011, Repo runs: Evidence from the tri-party repo market, *New York Federal Reserve Bank Staff Report* 506.

Cvitanić, Jakša, and Ioannis Karatzas, 1992, Convex duality in constrained portfolio optimization, *The Annals of Applied Probability* 2, 767–818.

Dang, Tri Vi, Gary Gorton, and Bengt Holmstrom, 2012, Ignorance, debt and financial crises, *Working paper*.

Diamond, Douglas W., and Philip H. Dybvig, 1983, Bank runs, deposit insurance, and liquidity, *The Journal of Political Economy* 91, pp. 401–419.

Duffie, Darrell, 1996, Special repo rates, *The Journal of Finance* 51, pp. 493–526.

——— , 2010, Presidential address: Asset price dynamics with slow-moving capital, *The Journal of Finance* 65, 1237–1267.

——— , Nicolae Garleanu, and Lasse Heje Pedersen, 2002, Securities lending, shorting, and pricing, *Journal of Financial Economics* 66, 307 – 339.

Fama, Eugene F., 1980, Banking in the theory of finance, *Journal of Monetary Economics* 6.

Flannery, Mark J., 1996, Financial crises, payment system problems, and discount window lending, *Journal of Money, Credit and Banking* 28, 804–824.

Fleming, Michael J., and Kenneth D. Garbade, 2003, The repurchase agreement refined: Gcf repo®, *Federal Reserve Bank of New York Current Issues in Economics and Finance* 9.

Friedman, Milton, 1960, *A Program for Monetary Stability* (Fordham University Press (New York)).

Geanakoplos, John, 1997, Promises, promises, in W.B. Arthur, S. Durlauf, and D. Lane, ed.: *The Economy as an Evolving Complex System II* (Addison-Wesley).

——— , 2010, The leverage cycle, in D. Acemoglu, K. Rogoff, and M. Woodford, ed.: *NBER Macroeconomic Annual 2009* (University of Chicago Press).

Gofman, Michael, 2011, A network-based analysis of over-the-counter markets, .

Goodhart, Charles, 1999, Myths about the lender of last resort, *International Finance* 2, 339–360.

Gorton, Gary, and Andrew Metrick, 2010a, Haircuts, *Federal Reserve Bank of St. Louis Review* 92.

——— , 2010b, Regulating the shadow banking system, *Brookings Papers on Economic Activity.*

——— , 2011, Securitized banking and the run on repo, *Journal of Financial Economics.*

——— , 2012, Who ran on repo?, Working Paper 18455 National Bureau of Economic Research.

Gorton, Gary, and Guillermo Ordonez, 2012, Collateral crises, Working Paper 17771 National Bureau of Economic Research.

Gromb, Denis, and Dimitri Vayanos, 2002, Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361 – 407.

Guvenen, Fatih, 2009, A parsimonious macroeconomic model for asset pricing, *Econometrica* 77, 1711–1750.

Hayek, F. A., 1976, *The Denationalization of Money: The Argument Refined* (The Institute of Economic Affairs).

He, Zhiguo, and Arvind Krishnamurthy, 2008, Intermediary asset pricing, Working Paper 14517 National Bureau of Economic Research.

He, Zhiguo, and Wei Xiong, 2012a, Dynamic debt runs, *Review of Financial Studies* 25.

———, 2012b, Rollover risk and credit risk, *The Journal of Finance* 67, 391–430.

Heston, Steven L., Mark Loewenstein, and Gregory A. Willard, 2007, Options and bubbles, *Review of Financial Studies* 20, 359–390.

Hordahl, Peter, and Michael R. King, 2008, Developments in repo markets during the financial turmoil, *BIS Quarterly, December 2008.*

Hugonnier, Julien, 2011, Rational asset pricing bubbles and portfolio constraints, .

Kiyotaki, Nobuhiro, and John Moore, 1997a, Credit chains, *Working paper.*

———, 1997b, Credit cycles, *The Journal of Political Economy* 105, pp. 211–248.

———, 2001, Evil is the root of all money, *Clarendon Lectures.*

———, 2012, Liquidity, business cycles, and monetary policy, Working Paper 17934 National Bureau of Economic Research.

Korinek, Anton, 2011, Systemic risk taking: Amplification effects, externalities and regulatory responses, .

Krishnamurthy, Arvind, 2002, The bond/old-bond spread, *Journal of Financial Economics* 66, 463 – 506.

———, Stefan Nagel, and Dmitry Orlov, 2012, Sizing up repo, Working Paper 17768 National Bureau of Economic Research.

Kyle, Albert S., and Wei Xiong, 2001, Contagion as a wealth effect, *The Journal of Finance* 56, 1401–1440.

Lorenzoni, Guido, 2008, Inefficient credit booms, *The Review of Economic Studies* 75, 809–833.

Mankiw, N.Gregory, and Stephen P. Zeldes, 1991, The consumption of stockholders and nonstockholders, *Journal of Financial Economics* 29, 97 – 112.

Martin, Antoine, David Skeie, and Ernst-Ludwig von Thadden, 2011, Repo runs, .

Mitchell, Mark, Lasse Heje Pedersen, and Todd Pulvino, 2007, Slow moving capital, *The American Economic Review* 97, pp. 215–220.

Monnet, Cyril, 2011, Rehypothecation, *Business Review* pp. 18–25.

———, and Borghan N. Narajabad, 2012, Why rent when you can buy? a theory of repurchase agreements, *Working paper.*

Parker, Jonathan A., and Annette Vissing-Jorgensen, 2010, The increase in income cyclicality of high-income households and its relation to the rise in top income shares, Working Paper 16577 National Bureau of Economic Research.

Pawasutipaisit, Anan, and Robert M. Townsend, 2011, Wealth accumulation and factors accounting for success, *Journal of Econometrics* 161, 56 – 81 <ce:title>Measurement with Theory</ce:title>.

Piketty, Thomas, 1997, The dynamics of the wealth distribution and the interest rate with credit rationing, *The Review of Economic Studies* 64, 173–189.

——— , and Emmanuel Saez, 2003, Income inequality in the united states, 1913–1998*, *Quarterly Journal of Economics* 118, 1–39.

Prieto, Rodolfo, 2010, Dynamic equilibrium with heterogeneous agents and risk constraints, .

Rampini, Adriano A., and S. Viswanathan, 2010, Collateral, risk management, and the distribution of debt capacity, *The Journal of Finance* 65, 2293–2322.

Repullo, Rafael, 2005, Liquidity, risk-taking, and the lender of last resort, *International Journal of Central Banking* 1, 47–80.

Rochet, Jean-Charles, and Jean Tirole, 1996, Interbank lending and systemic risk, *Journal of Money, Credit and Banking* 28.

Shin, Hyun Song, 2010, *Risk and Liquidity* (Oxford University Press).

Singh, Manmohan, 2011, Velocity of pledged collateral, *IMF Working Paper*.

——— , and James Aitken, 2009, Deleveraging after lehman - evidence from reduced rehypothecation, *IMF Working Paper*.

——— , 2010, The (sizable) role of rehypothecation in the shadow banking system, *IMF Working Paper*.

Smith, Josephine, 2012, Collateral value risk in the repo market, *Working paper*.

Stein, Jeremy C., 2012, Monetary policy as financial stability regulation, *The Quarterly Journal of Economics* 127, 57–95.

Thornton, H., 1802, *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain* (London: Hatchard).

Townsend, Robert, and Neil Wallace, 1987, *Circulating Private Debt: An Example with a Coordination Problem* (University of Minnesota Press).

Vayanos, Dimitri, and Pierre-Olivier Weill, 2008, A search-based theory of the on-the-run phenomenon, *The Journal of Finance* 63, 1361–1398.

Vissing-Jorgensen, Annette, 2003, Perspectives on behavioral finance: Does "irrationality" disappear with wealth? evidence from expectations and actions, *NBER Macroeconomics Annual* 18, 139–194.

——— , 2009, Who bears aggregate fluctuations and how?, *American Economic Review* 99, 399–405.

Wachter, Jessica A., and Motohiro Yogo, 2010, Why do household portfolio shares rise in wealth?, *Review of Financial Studies* 23, 3929–3965.

Wang, Jiang, 1996, The term structure of interest rates in a pure exchange economy with heterogeneous investors, *Journal of Financial Economics* 41, 75 – 110.

Zawadowski, Adam, 2013, Entangled financial systems, *Review of Financial Studies*.