Optimal Capital Regulation with Two Banking Sectors

Taejin Kim and Vishal Mangla *

January 16, 2013

Job Market Paper

Abstract

We present the case for procyclical capital requirement policy – lower requirement during booms and higher requirement during recessions – as opposed to the generally accepted countercyclical requirement in the current literature. Our argument is based on the fact that banks shift their capital into and out of the regulation purview depending upon the business cycle and the severity of the regulation. The banks trade off the benefit of being regulated – cheaper funding/insurance – with the cost – restriction on the portfolio risk. Tightening the capital requirement during a boom (as in a countercyclical policy) forces the banks to move to the shadow banking sector where they take too much risk. Therefore, the policy aimed at controlling the risk banks take during booms should incentivize the banks to be regulated by relaxing the capital requirement. These forces are reversed during recessions. The policy specifies the level of capital requirement as a function of the relative size of the two banking sectors. Under this specification, the right mix of the two sectors is achieved even when, in equilibrium, the banks are indifferent between being regulated and unregulated. We show that the regulation policy is robust to estimation errors in the model parameters in that small errors lead to welfare loss that is an order smaller.

Keywords: Capital Regulation, Shadow Banking, Regulatory Arbitrage, Optimal Debt Contract.

*Northwestern University. We are deeply indebted to Arvind Krishnamurthy for his guidance. We are grateful for helpful comments from Mirza Afrasiab (discussant), Snehal Banerjee, Alexander Bleck (discussant), Michael Fishman, Craig Furfine, Pedro Gete, Kathleen Hagerty, Guido Lorenzoni, Robert McDonald and Zhenyu Wang (discussant). We also thank seminar participants at Cleveland Fed, Econometrics Society NASM 2012, EFA 2012, Northwestern University and WFA 2012. All errors are our own. E-mails: {taejin-kim, v-mangla}@northwestern.edu.

1 Introduction

Financial intermediaries shift their capital across the regulated and the unregulated banking sectors depending upon the business cycle and the severity of regulation.[1] In making this decision, they compare the benefit of being regulated – access to cheaper funding due to the government guarantees – with the cost – restriction on the portfolio risk. During booms, as the investment opportunities improve, the benefit of being regulated shrinks because the funding becomes easier whereas the cost of being regulated grows due to the increasing shadow value of the portfolio constraint. Therefore, when the level of the regulation is held fixed, financial intermediaries are attracted towards the unregulated sector during booms. These forces reverse during recessions.

We saw an evidence of this regulatory arbitrage by financial intermediaries before and during the crisis of 2007-2009. Before the crisis, the unregulated/shadow banking sector experienced a spectacular growth (Coval, Jurek and Stafford (2009), Gorton and Metrick (2010)). Figure 1 illustrates that the size of the shadow banking sector doubled from $10 trillion in the year 2000 to $20 trillion at its peak in March 2008. An important component of this growth was the widespread adoption of off-balance sheet investment vehicles such as SIVs and SPVs by means of which large commercial banks shifted their investments out of the regulation purview (Acharya, Schnabl, and Suarez (2010), Gorton and Metrick (2010)). On the flip side, during the crisis, Goldman Sachs and Morgan Stanley partially shifted their capital into the regulation purview by acquiring the status of bank holding company (NY Times (2009)). More generally, a number of hitherto unregulated entities of the intermediary sector accepted the government support. Figure 2 captures the cyclic behavior of the growth of the relative size of the unregulated banking sector.

The current literature on capital requirement regulation fails to consider the strategic capital flows between the two banking sectors, and focuses on controlling the risk assumed only by the regulated banking sector over the business cycle. However, it was the risk taking in the shadow banking sector that led to the recent financial crisis. Thus, even though shadow banks are outside the direct control of regulator by definition, a desirable policy should take into account the risk

[1]For our purpose, unregulated/shadow banks include structured investment vehicles (SIVs) and special purpose vehicles (SPVs) which intermediate credit through securitization and secured funding techniques such as asset-backed commercial paper (ABCP), asset-backed securities (ABS) and collateralized debt obligations (CDOs). However, finance companies, credit hedge funds, money market mutual funds, securities lenders and the government-sponsored enterprises (GSEs) are also examples of shadow banks (Pozsar et al. (2010)).

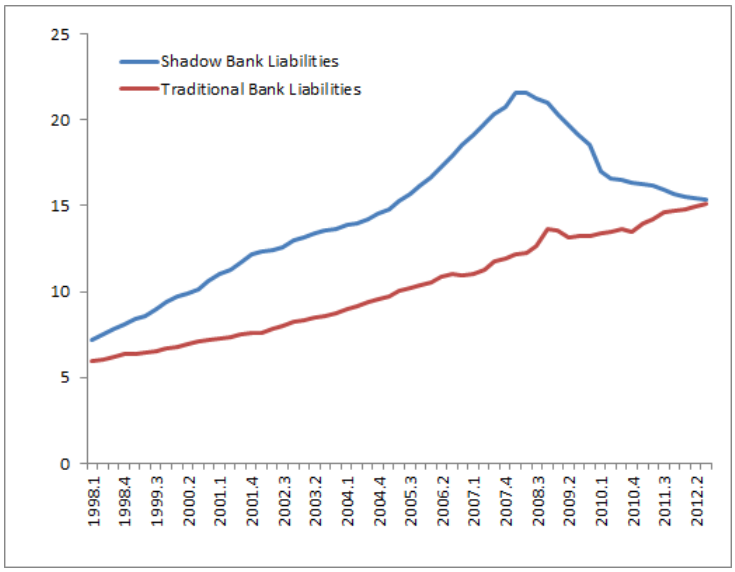


Figure 1: Shadow Bank Liabilities vs. Traditional Bank Liabilities, $ *trillion*.
Source: Flow of Funds Accounts of the United States as of 2012:Q3 (FRB) and FRBNY.
Shadow bank liabilities correspond to securitization activity and short term money market transactions that are not backstopped by deposit insurance. The details behind the construction of this plot are provided in Pozsar et al. (2010).

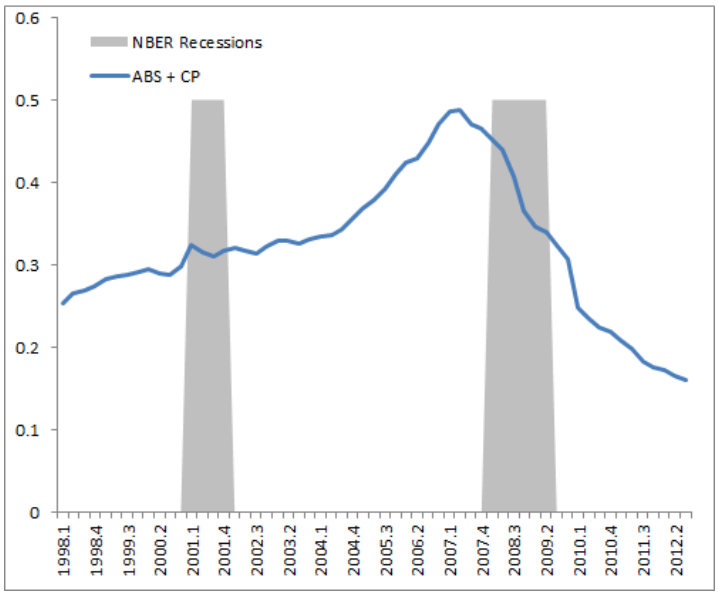


Figure 2: Cyclic behavior of Asset Backed Securities (ABS) plus unenhanced Commercial Paper (CP).
Source: Flow of Funds Accounts of the United States as of 2012:Q3 (FRB) and NBER.
We plot the ratio of the liabilities of the ABS issuers (Line 11 of Table L.124 in Flow of Funds) plus CP without explicit official guarantee (sum of Lines 6, 10, 12 and 13 of Table L.208) and the liabilities of the traditional banking sector (Line 19 of Table L.109).

taken by both the regulated and the unregulated banking sectors. This is what we do in this paper. We find the capital regulation policy that controls the aggregate risk in the economy. The most interesting outcome of our analysis is the procyclical nature of the optimal capital requirement policy – capital requirement is relaxed during booms and tight during recessions. This policy is the exact opposite of the suggestion in the current literature which argues for countercyclical requirements (example Kashyap and Stein (2004)).

The current literature proposes a higher capital requirement during good times which can act as a buffer against losses during bad times. However, this advice ignores the fact that banks that are averse to raising capital will simply switch their types from regulated to unregulated (move their investment out of the regulator's purview as they did before the crisis). This will in turn result in too big a shadow banking sector and the corresponding high systematic risk. To limit the risk in good times, we propose that the government actually requires a lower capital ratio which will entice the banks to become regulated and thus will not take too much risk (since the regulated banks need to hold capital commensurate to the risks of their investments). Similarly, the current view proposes lower capital requirement during bad times so that banks are not forced to cut credit and/or fire-sell their assets in order to maintain their capital ratio. However, too low capital requirement means that all banks would prefer to be regulated since they receive government's funding support at a low cost. So, there would be too much capital in the regulated banking sector and the economy-wide investment profile would be overly conservative. To limit this excessive conservatism during bad times, we propose that the government actually requires a higher capital ratio so that some banks find it unprofitable to be regulated. This will balance the systematic risk during bad times.

It should be noted that, as in the numerous studies evaluating the effectiveness of the Basel II standards, our model generates procyclical risk-taking when the capital requirement is time-invariant. In tune with the concern of the current capital regulation debates, we seek to mitigate the excessive procyclicality of the economy through the capital regulation. Since the role of the regulation is to introduce a force counter to the procyclical economic fluctuation, the counter-cyclical capital requirement proposed in the literature is a natural suggestion. However, this conclusion implicitly assumes that the banks cannot escape regulation. One of the main points

of this paper is that, if we take into consideration the strategic behavior of banks, the capital requirement policy is not as obviously countercyclical as it made out to be in the current literature. To clearly derive our procyclical policy result, we ignore the traditional forces that generate countercyclical policy. For example, banks may face significant costs in switching their types, or the two types of banks may not have access to the same assets. However, in Section 9 we consider an extension of our model where even when we introduce one of the traditional forces – changing bank capital over business cycle – the procyclical regulation result survives.

In our economy, there are two projects available for investment: a riskless project and a risky project. There are three agents: banks, investors and a regulator. While the investors are risk-averse, the banks are risk-neutral. The investors themselves have no access to the risky project and obtain risk exposure by investing in the banks' securities. The banks have access to both the projects. If they are unregulated, the banks invest only in the risky project. The benefit of being regulated is cheaper financing from the investors due to deposit insurance provided by the regulator (we assume zero fee for deposit insurance). The cost of being regulated is a regulator imposed capital requirement which takes the following rule: the risk of a bank's asset portfolio return is upper bounded by a scalar λ times the bank's equity capital. The multiple λ is the policy instrument of the regulator and the same for all the banks. Each bank compares the benefit and the cost of being regulated in deciding its type. Under the absence of any regulation, all investment flows into the risky project (since banks are risk-neutral) and the social welfare is below the first-best.[2] The regulator's objective is to maximize the social welfare which, in this setup, amounts to obtaining the right amount of the aggregate investment in the risky project. We show that the optimal level of capital requirement λ varies with the underlying economic fundamentals in a procyclical manner.

In our model, each bank is indifferent between being regulated and unregulated in equilibrium. This implies that we cannot determine the capital sizes of the two banking sectors only from the equilibrium condition. Also, for the same amount of equity capital, an unregulated bank invests more in the risky project than a regulated bank. It follows that the aggregate investment in the risky project (or 'risk exposure') is not determined in equilibrium. We resolve this indeterminacy

[2]We provide the rationale for the necessity of regulatory intervention in more detail in section 2.4.

issue by making the level of capital requirement λ directly depend on a macro variable that is related to the aggregate risk exposure and is an observable to the regulator. In our model, the relative size of the two banking sectors pins down the aggregate risk exposure. We assume that the regulator is able to estimate the capital size of the unregulated banking sector. The optimal policy specifies λ as a function of the relative sector size realized at a given time such that whenever the relative sector size is suboptimal, λ, as specified by the mapping, incentivizes the banks to switch to the right sector. To see how this policy works, suppose there are too many unregulated banks in the economy. Then, under our policy, the capital requirement would be too relaxed, making it profitable for the unregulated banks to convert to regulated banks. But this would reduce the relative size of the unregulated banking sector and consequently, the capital requirement would tighten. The unregulated banks will continue to move to the regulated sector until the capital requirement is tight enough to make them indifferent between the two sectors. We pick our schedule of the capital requirement such that the banks' indifference is achieved exactly when the relative sector size is optimal (i.e., the corresponding risk exposure is optimal).

The regulation policy proposed in this paper has several desirable features expected of a macro policy. First, the policy is macroprudential in that the objective of the policy is to control the aggregate risk assumed by the financial intermediary sector (rather than focusing on the health of each individual bank). Second, the policy is market-based. Due to the feedback loop that exists between the capital requirement and the relative size of the two banking sectors, once the policy schedule has been announced ('etched in stone'), the market self-adjusts to attain the optimal risk exposure due to the mechanism described above.[3] Third, this feedback mechanism also ensures that the economy stays on the equilibrium path as the underlying fundamentals change, making the policy robust to the business cycle fluctuations. Fourth, we show that the policy is robust to estimation errors in the model parameters in that small errors lead to welfare loss that is an order smaller.

There are three main empirical predictions of our model: (i) the size of the shadow banking sector relative to that of the commercial banking sector is procyclical when the capital requirement

[3] A feedback loop between the policy instrument and the macro variable being targeted is also a feature of Taylor rule (Taylor (1993)) in monetary economics where the federal funds rate (the policy instrument) is allowed to vary with the actual inflation rate and the unemployment rate (the macro variables) according to a pre-specified rule in order to arrive at their target values.

is held fixed; (ii) the leverage of the shadow banking sector is procyclical; and (iii) the shadow banks offer procyclical interest rate to their investors. The first two predictions find support in data. For example, Adrian and Shin (2010) show that marked-to-market leverage of financial intermediaries – in particular, broker-dealers – is strongly procyclical, because those firms actively engage in the management of balance sheets, responding to the changes in economic conditions. The terms of debt contract in this paper clearly demonstrate the channel through which leverage responds to the business fluctuation in a procyclical manner.

The underlying spirit of this paper is very different from that of Gorton and Metrick (2010), Hanson, Kashyap and Stein (2011) and Dodd-Frank Act 2010 where the focus is to identify all entities engaging in financial intermediation and impose regulation on them. However, the history of the financial intermediation sector makes us believe that banks have always been able to find means to get around regulation (for an example see Acharya and Richardson (2009)). If this pattern continues to hold, the best regulators can do is to implicitly influence (and not overtake) the shadow banking sector by designing regulation as an incentive scheme.[4] Regulatory arbitrage as a result of the capital requirement has also been emphasized by Goodhart et al. (2011). Like us, they introduce a 'shadow banking' sector that provides intermediation along with a 'banking' sector. In their model, the disadvantage of a big shadow banking sector is the exacerbation of the fire sale problem during a bust. However, rather than proposing an optimal capital regulation, their focus is on contrasting the relative performance of five different policy instruments commonly advocated (that is, choose the 'right policy instrument' rather than the 'right level' of a given policy instrument).

2 Setup

We model a real economy. There are two dates $t = 0, 1$.[5] All agents make their decisions at $t = 0$ and investments pay off at $t = 1$. There is no information asymmetry between investors and banks. The banks can invest in risky project but the investors cannot.

[4]Except for some extreme values of the model parameters, we find that it is always optimal for the government to have a strictly positive mass of the shadow banking sector.

[5]We will introduce the possibility of changing economic fundamentals in $t \in (0, 1)$ later.

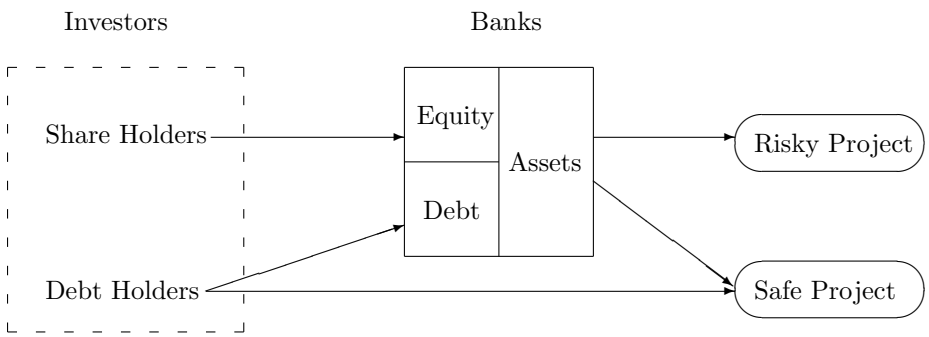


Figure 3: Structure of the Economy

2.1 Investment Projects

There are two investment projects: riskless and risky. The riskless project acts as a safe storage – it generates unity cash flow at $t = 1$ for every unit of investment at $t = 0$. The risky project (best thought of as a risky production technology) generates a stochastic cashflow $\tilde{R}$ at $t = 1$ for every unit of investment at $t = 0$. The projects are linearly scalable.

The distribution of $\tilde{R}$ is F, with mean μ (with $\mu > 1$) and standard deviation σ. The density f is continuously differentiable. F is completely determined by μ and σ, denoted by $F^{\mu,\sigma}$. Moreover, all $F^{\mu,\sigma}$ belong to one location-scale family, like the Normal distribution. Hence we write $f^{\mu,\sigma}(\tilde{R}) = \frac{1}{\sigma} f^{0,1}(\frac{\tilde{R}-\mu}{\sigma})$. We drop the superscripts μ and σ when there is no confusion.

2.2 Banks

There is a continuum of banks of measure one with risk-neutral bankers. Bank j (where $j \in [0, 1]$) has equity capital k_j. Except for the size of their equity capital, the banks are identical. They have the same information and investment opportunity set. We assume no heterogeneity in their skills or effort levels. The aggregate equity capital of the banking sector is $K \equiv \int_0^1 k_j dj$.

Since the banks are risk-neutral, they invest only in the risky project in the absence of any constraint. Risk-neutrality greatly simplifies the banks' portfolio choice problem but is not essential for our results. However, we still require that the banks are less risk averse than the investors. Since the investors have no direct access to the risky project, this misalignment of preferences

creates friction in the banks-investors relationship – banks make investments that investors find too risky.

If a bank with equity capital k raises debt d from the investors and invests α (with $0 \leq \alpha \leq 1$) fraction of its funds in the risky project, the (gross) portfolio return realized at $t = 1$ is

$$\alpha(k + d)\tilde{R} + (1 - \alpha)(k + d) \tag{1}$$

Each bank can choose to be one of the two types: (i) unregulated bank (*shadow bank, SB*) or (ii) regulated bank (*commercial bank, CB*).[6] The SBs face no portfolio constraint and set $\alpha = 1$ but they have to pay the market interest rate on the debt they raise – the shadow banks' leverage and the interest rate are endogenously determined in our model. In contrast, the debt raised by the CBs is insured by the regulator and thus the CBs pay the risk-free interest rate of unity (gross). The CBs receive the debt insurance for free.[7] They set α equal to the maximum value allowed under the regulation (details to follow). We assume that a bank's type does not affect its accessibility to the projects.[8]

Examples of commercial banks are Bank of America and Citigroup that take deposits from investors. These deposits are insured by the government. The shadow banking sector includes investment banks (Lehman, Goldman Sachs (before crisis) etc.), finance companies, asset-backed commercial paper (ABCP) conduits, limited-purpose finance companies, structured investment vehicles, credit hedge funds, money market mutual funds, and securities lenders.

2.3 Investors

There are two continuums of investors – a measure W of investors each endowed with a unit of good/cash (cash holders) and a measure K of investors each endowed with a share of the aggregate

[6]In this model, a bank can be of only one type. This may seem different from a real world commercial bank setting up a special purpose vehicle (SPV) to move capital outside the regulation purview. This distinction is not important from the regulator's perspective; what matters is the ratio of unregulated to regulated bank capital in the economy.

[7]Charging the commercial banks insurance premium that depends on their investment portfolio could be one way to restrict the amount of risk they take. However, in tune with the policy practice, we instead focus on capital regulation to restrict the commercial banks' investments.

[8]This assumption has become more reasonable recently, mainly due to the securitization technology and the rapid growth of its products. However, the traditional loans are harder to be substituted by the unregulated sector because relational effects and soft information are valuable. In other words, with traditional loans, banks face higher level of friction in moving their capital. Aiyar, Calomiris and Wieladek (2012) focus on this issue.

banking sector (share holders). The investors have the same risk preference represented by a Bernoulli function $u(\cdot)$ with $u(1) = 0$, $u' > 0$, $u'' < 0$ and satisfying the integrability condition $\int_0^\infty |u(\tilde{R})| dF < \infty$. The investors do not have access to a market to trade their shares and cash. Also, they do not have a direct access to the risky project. No individual share holder is capable of negotiating with the banks over their portfolio choice.[9] An individual cash holder can invest her unit of good either in a bank's debt (and thus become a debt holder) or in safe storage (and thus stay a cash holder) but cannot split between the two.[10]

After the cash holders have lent to the banks but before the investments pay off, all the investors in the economy – equity holders, debt holders and cash holders – encounter an unanticipated pooling opportunity. The investors have the option of collectively pooling their investments and receive the fraction $1/(W+K)$ of the aggregate cashflow at $t = 1$. We will see that under mild restrictions on the model parameters, it is individually rational for the investors to collectively pool their assets.

Our modeling device of the unanticipated pooling opportunity is a particular way to make the market equilibrium (the equilibrium that obtains when all the banks are shadow banks) Pareto inefficient.[11] We take as given the premise that the aggregate risk in the economy matters for the welfare while the investors cannot attain the optimal level of aggregate risk through their investment in the banks on their own. In particular, the economy with no intervention is exposed to excessive fluctuations, leading to a need for regulatory intervention. The modeling device in this section enables us to build a model that is simple but still satisfies the premise by stopping the investors from getting around excessive risk taking of banks through a channel artificially generated by the simple setting. When we discuss the form of regulator's intervention in the next subsection, it will be become clear why the particular form of friction that delivers the inefficiency is not critical for our qualitative results.[12]

[9] This could be due to some managerial frictions that prevent the investors from punishing the bankers when the bankers fail to carry out the investment plan agreed upon with the investors.

[10] When a bank's debt is risky, a portfolio choice problem naturally arises for the cash holders. Restricting the cash holders' portfolio choice lets us avoid the associated computational complexity but is not essential for the inefficiency present in our setup (discussed below).

[11] Here, the social welfare is the sum over all the investors of the expected utility from consumption. The details are in the next subsection.

[12] Another plausible friction that creates the inefficiency is to have the share holders and the cash holders belong to the same household but they make independent investment decisions.

2.4 Regulator

In our model, when all the banks are shadow banks, the aggregate cashflow is too risky; it is riskier than the first best in which the investors could invest directly in the risky project. The inefficiency arises because the banks are able to do both – (i) obtain too much funds from the debt holders, and (ii) invest freely (i.e., no portfolio constraint). The inefficiency vanishes if any one of the two conditions is not satisfied. Introducing a regulator equipped with the tool to impose portfolio restriction on the commercial banks breaks the second condition and achieves Pareto improvement.

2.4.1 Objective

Since both the safe storage and the commercial banks' debt deliver the risk-free return of unity, the two are identical from the investors' perspective. Therefore, we impose for simplicity that the cash holders who wish to invest in the safe storage simply hold the commercial banks' debt. Under this convention, the total assets held by all the banks is $W + K$, with K in the form of equity and W in the form of debt. The portion of these aggregate assets that is held by the shadow banks is fully invested in the risky project whereas the portion that is held by the commercial banks is split between the risky project and the safe project.

The objective of the regulator is to maximize the investors' welfare from the net cashflows they receive. The investors receive the following cash inflows on aggregate: the debt and the equity payments from all the banks, and the insurance payouts by the regulator when the commercial banks default. On the other hand, the investors have to fund the regulator's insurance program (in the form of some tax that we do not model explicitly). The net of these cash outflows and inflows to the investors is simply the aggregate cashflow generated from all the banks' investments.[13] Denote this net cashflow by $\tilde{Y}$. The regulator's objective is to choose the level of capital requirement so as to maximize the social welfare:

$$(W + K)\ \mathbb{E}\left[u\left(\frac{\tilde{Y}}{W + K}\right)\right] \tag{2}$$

[13]Eventually, the economy cannot consume more than the total return from investments.

Suppose, at the aggregate level, α fraction of the banks' total assets $W + K$ flows into the risky project. Then, the aggregate cashflow generated from the banks' investments is simply

$$\tilde{Y} = (W + K)[1 + \alpha(\tilde{R} - 1)] \tag{3}$$

under which the social welfare (2) has the form:

$$(W + K)\ \mathbb{E}[u(1 + \alpha(\tilde{R} - 1))]$$

Therefore, the regulator's objective boils down to implementing the α that maximizes $\mathbb{E}[u(1 + \alpha(\tilde{R} - 1))]$. The solution α^* to this problem is some function $\phi(\mu, \sigma)$ with the property $\phi_\mu > 0$ and $\phi_\sigma < 0$.

2.4.2 Regulation

If a bank chooses to be regulated, it receives deposit insurance for free. In return, the regulator stipulates that the bank's portfolio satisfies the following condition:

$$\text{standard deviation of the bank's portfolio return} \leq \lambda k, \tag{4}$$

where $\lambda > 0$ is the policy instrument (capital requirement) and k is bank's equity capital. This regulation puts a restriction on how much risk a bank can take given its equity capital. This specification is consistent with the Basel II accord – it stipulates the minimum amount of risk-based capital (note higher is the standard deviation of bank's portfolio return, higher is its credit risk or probability of default). However, unlike Basel II, we let λ vary with the fundamentals (μ and σ).

If a bank holds the portfolio (1), the regulation (4) implies that the regulated bank can invest in the risky project up to the limit

$$\alpha = \frac{\lambda k}{\sigma(k + d)}, \tag{5}$$

where d is the amount of debt raised by the bank. The bank's leverage for its investment in the

risky project ('risk leverage') is $\alpha(k+d)/k$ which under the above regulation is λ/σ. So, we see that the regulation (4) amounts to a limit on a commercial bank's risk leverage. Obviously, if a bank invests its deposits in the riskless project, the regulator does not need to regulate the bank's leverage because there is no chance of default; the regulator is only concerned about the fraction of bank's investment in the *risky* project.

3 Debt Contracts

We now investigate how the terms of debt contract – interest rate and leverage – are determined for the shadow banks and the commercial banks. Asset-backed commercial paper (ABCP) and deposits are examples of debt security offered by shadow and commercial banks respectively. We will assume that the banks possess the whole bargaining power and extract the total rent from their debt holders – the terms of the debt contract are such that the debt holders are driven to their outside option of simply storing the cash. The degree of bargaining power of each counterparty is not crucial for our model, but what we require is that the terms of contract are decided by a take-it-or-leave-it bargaining and the banks move first.

3.1 Shadow Banks' Debt Contract

A shadow bank with equity capital k and leverage $l(k)$ (asset to equity ratio)[14] offers a take-it-or-leave-it interest rate $r(k)$ to a potential debt holder. An individual cash holder accepts the offer and lends her unit of cash to the bank if it is individually rational for her to do so. Since the riskiness of the bank relies on the bank's leverage, the individual rationality depends on the bank's leverage too. Since each debt holder provides only one unit of cash to the bank, the leverage $l(k)$ necessary to determine the individual rationality is an expected value of the debt holder. Under rational expectation equilibrium, the mass of cash holders that lend to this bank is such that the bank's leverage is actually $l(k)$. This means that the bank has $k[l(k)-1]$ measure of debt holders.

[14] If a bank has equity capital k and raises debt d, its leverage is $l \equiv (k+d)/k$.

Then, equilibrium $r(\cdot)$ and $l(\cdot)$ are the solution to the optimization problem:

$$\max_{r(\cdot)>1, l(\cdot)>1} \mathbb{E} \max \left\{ kl(k)\tilde{R} - k[l(k)-1]r(k), 0 \right\} \tag{6}$$

$$\text{s.t. } \mathbb{E}u\left(\min \left\{ r(k), \frac{l(k)}{l(k)-1}\tilde{R} \right\} \right) = u(1) \tag{7}$$

To see this, notice that the bank's total asset size is $kl(k)$ and the total return at $t = 1$ is $kl(k)\tilde{R}$. At $t = 1$, the bank owes $r(k)$ to each of its $k[l(k) - 1]$ debt holders. If the bank is unable to service its debt, it defaults, earns zero itself, and distributes the whole return equally among its debt holders in which case each debt holder receives $l(k)\tilde{R}/[l(k) - 1]$.

The particular structure of the problem greatly simplifies the analysis. Notice that once we factor out k from the objective function (6) (to obtain expected profit per unit capital), k enters the optimization problem only through $r(k)$ and $l(k)$. This means the problem is homogeneous in bank capital k. This observation delivers the following simplification:

Lemma 1. *The functions $r(\cdot)$ and $l(\cdot)$ are constant functions.*

This result ensures that the shadow banks' expected profit is linear in its equity capital k. Since the size of equity capital is the only heterogeneity of banks, this linearity guarantees that all the banks are identical in choosing their contract terms.

In the following, we will show that the solution (r, l) to the contracting problem stated above is unique. But before we are able to do so, we have to characterize the banks' profit function (6) and the investors' individual rationality (IR) constraint (7).

Henceforth, for mathematical convenience, we will work with a transformation of l called asset to liability ratio s and defined as[15]

$$s \equiv \frac{l}{l-1} \tag{8}$$

[15] If a bank has equity capital k and raises debt d, its asset to liability ratio is $s \equiv (k + d)/d$. Morever, $l = s/(s - 1)$.

3.1.1 Debt Holders' IR Constraint

Given that r and l are independent of k, condition (7) delivers a relation between l and r, or equivalently between s and r, denoted by $s^u(r)$. We can interpret that $1/(s^u(r) - 1)$ is the investors' supply curve – the amount of funds the debt holders are willing to supply to a bank with unit capital when it offers interest rate r.

Lemma 2. *The function $s^u(r)$ is decreasing and convex.*

A decreasing $s^u(r)$ means that the shadow banks can obtain higher leverage when they offer higher interest rate. The next lemma identifies the condition that prevents the shadow banks from obtaining infinite leverage, that is, $s^u > 1$.

Lemma 3. *If $\mathbb{E}u(\tilde{R}) < u(1)$, s^u is bounded away from 1. If $\mathbb{E}u(\tilde{R}) > u(1)$, there exists $\overline{r}$ such that $s^u(r) = 1$ for all $r > \overline{r}$.*

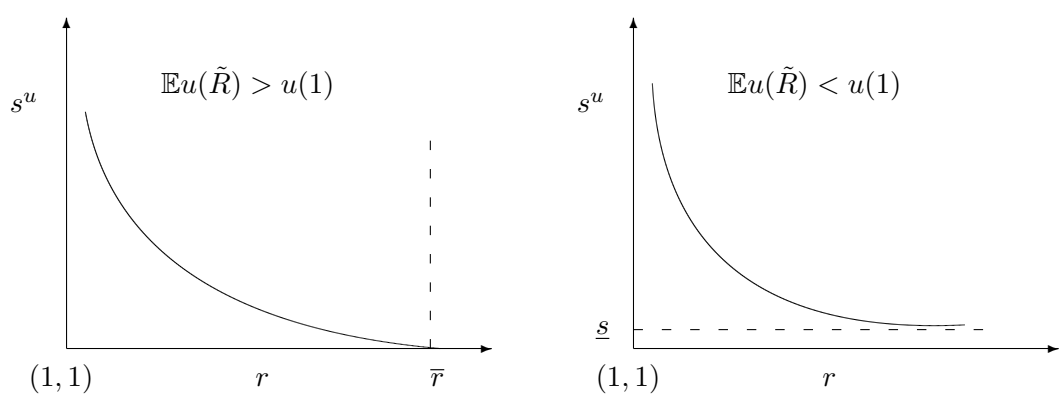


Figure 4: Finite Supply of Debt

If $\mathbb{E}u(\tilde{R}) > u(1)$, the debt holders are willing to lend infinite amount of funds for a sufficiently large but finite interest rate ($r \geq \overline{r}$). If $\mathbb{E}u(\tilde{R}) < u(1)$, the supply has an upper bound for all r. Figure 4 depicts these two cases. From here on, we assume the condition that ensures finite supply of debt.

Assumption 1. *(Finite Supply of Debt)*

$$\mathbb{E}u(\tilde{R}) < u(1)$$

This condition ensures that at least one of the following is true: (i) the investors are 'sufficiently' risk averse or, (ii) the risky project return is not 'too good'. Or, to quantify these adjectives, a cash holder is better off storing her unit of good than investing it in the risky project if she has access to the risky project.

3.1.2 Shadow Banks' Profit Function

It is convenient to define the following function related to the banks' expected profit:

$$\Pi(x, y; \mu, \sigma) \equiv \frac{y}{y-1}\left(\mathbb{E}\left[\tilde{R}\middle|\tilde{R} > \frac{x}{y}\right] - \frac{x}{y}\right)\left(1 - F\left(\frac{x}{y}\right)\right)$$

As we will see shortly, this function computes the per-unit-equity profit of a bank that has the asset to liability ratio y and pays the interest rate x.

Given Lemma 1 and the transformation (8), the shadow banks' expected profit (6) is rewritten as

$$\begin{aligned}\Pi^{SB}(k) &= \mathbb{E}\max\left\{\frac{ks}{s-1}\tilde{R} - \frac{kr}{s-1}, 0\right\} \\ &= \frac{ks}{s-1}\left(\mathbb{E}\left[\tilde{R}\middle|\tilde{R} > \frac{r}{s}\right] - \frac{r}{s}\right)\left(1 - F\left(\frac{r}{s}\right)\right) \\ &= k\Pi(r, s)\end{aligned} \tag{9}$$

The following lemma states two important properties of the marginal rate of substitution of Π, $MRS^{\Pi} \equiv \frac{\partial\Pi/\partial r}{\partial\Pi/\partial s}$.

Lemma 4. *The marginal rate of substitution of the iso-profit function is positive and decreasing*

in r.

That is, the banks' iso-profit curves are decreasing and convex in the plane with r on the horizontal axis and s on the vertical axis.

3.1.3 Solution to the Contracting Problem

With the characterizations of the debt holders' IR constraint and the shadow banks' profit function, we turn to the contracting problem. Solving the original contracting problem boils down to finding r that solves

$$\max_{r>1} \Pi(r, s^u(r))$$

Accordingly, define a continuous function

$$\Pi(r) \equiv \Pi(r, s^u(r))$$

Let MRS^U denote the marginal rate of substitution of the debt holders' IR curve (it is equal to $-ds^u(r)/dr$). The following lemma characterizes the solution of the above problem.

Lemma 5. *For any Bernoulli function u, the following programs have the same solution(s) for r.*

1. $\Pi'(r) = 0$
2. $MRS^U = MRS^\Pi$ *and* $s = s^u(r)$
3. $\Pi(r) = r \int_0^r \left(\frac{u'(x)x}{u'(r)r} - 1\right) f\left(\frac{x}{s^u(r)}\right) \frac{1}{s^u(r)} dx + r \equiv J(r)$

The function $J(\cdot)$ lacks an intuitive interpretation, but we can easily show that, in the CRRA(γ) case, it takes a simple form, $J(r) = r^\gamma$.

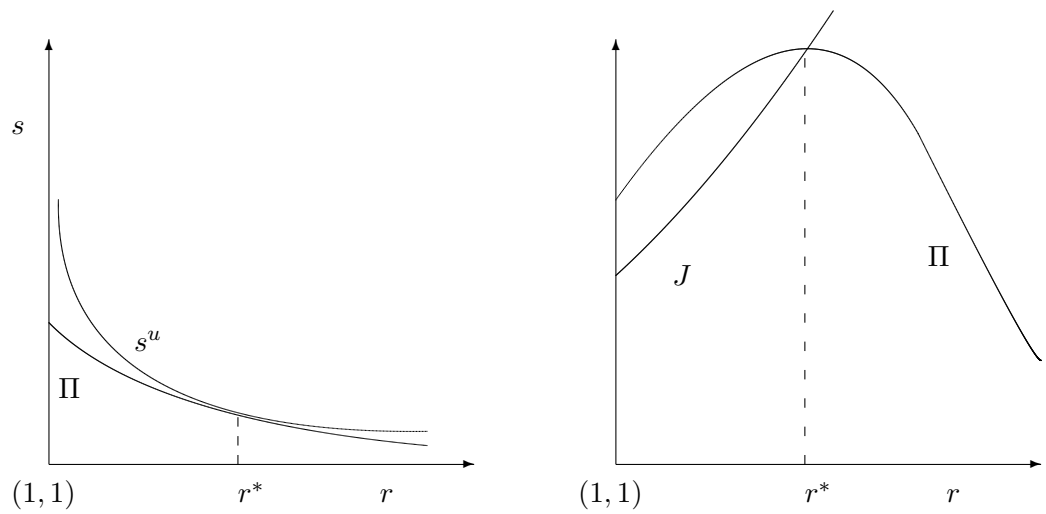


Figure 5: Solution to the Contracting Problem

Figure 5 displays how the contract term, r^*, is determined. The left figure shows that the two marginal rates of substitution coincide, the second statement in Lemma 5. The right figure illustrates the third statement in Lemma 5 that the function $J(r)$ cuts $\Pi(r)$ at the solution. The next lemma proves that the function $J(r)$ is increasing in r, as in Figure 5.

Lemma 6. *$J'(r) > 0$, $\lim_{r\to 1} J(r) = 1$, and $\lim_{r\to\infty} J(r) = \infty$.*

Lemma 6 plays a crucial role in proving the following proposition.

Proposition 1. *The solution (r^*, s^*) to the contracting problem exists and is unique.*

Throughout our analysis, we assume that the cash holders' wealth is large relative to the banking sector equity capital: $K/W < s^* - 1$. This condition ensures that the shadow banks face no constraint in raising debt due to limited cash holders' wealth. When all the banks are shadow banks, the total amount of debt they raise is $K/(s^* - 1)$ (in this case, the excess cash holders' wealth is kept in the safe storage).

3.2 Commercial Banks' Debt Contract

This contract is simple. Due to the deposit insurance, the commercial banks' debt is risk-free and thus the interest rate offered by the commercial banks is 1 (the risk-free rate). Due to this, any individual commercial bank's leverage is indeterminate; the commercial banks' debt security simply acts as a store of money for the cash holders in our model. We will see shortly that the expected profit of a commercial bank is independent of the debt it raises. However, the aggregate leverage of the commercial banking sector is pinned down by the market clearing condition in the cash holders' wealth – the cash holders who do not become the shadow banks' debt holder become the commercial banks' debt holder.

Now let us consider what happens when the investors encounter the unanticipated pooling opportunity after they have made their investments. We show that it is individually rational for the investors to collectively pool their investments.[16] By doing so, they earn an aggregate cashflow that is the sum of the cashflow generated by the entire banking sector, $\tilde{Y} = (W+K)[1+\alpha(\tilde{R}-1)]$ (from (3)), and the deposit insurance payouts by the regulator. Therefore, if the regulator is able to implement the optimal $\alpha = \phi(\mu, \sigma)$, the expected utility of each investor is strictly more than $u(1)$ (since $\phi(\mu, \sigma) > 0$) if all the investors pool. However, if they were not to pool, the debt holders of the shadow banks earn an expected utility of $u(1)$ due to the IR constraint (7) and the debt holders of the commercial banks simply earn $u(1)$. Moreover, since the share holders have more than 100% exposure to the risky project return, Assumption 1 implies that their expected utility is below $u(1)$. Therefore, each investor is better off by pooling her investment with the other investors.

4 Banks' Choice: Commercial vs. Shadow

Given the debt contracts of the previous section, we now consider the decision of a bank to become regulated (commercial). If a commercial bank with equity capital k raises debt d, the fraction it invests in the risky project is given by (5). Therefore, the expected profit of the commercial bank

[16] We impose that the investors can either pool collectively or not pool at all. That is, we do not allow any partial pooling of the investors.

is

$$
\begin{aligned}
\Pi^{CB}(k) &= \mathbb{E}\max\left\{\left(\frac{\lambda}{\sigma}k\tilde{R} + k + d - \frac{\lambda}{\sigma}k\right) - d, 0\right\} \\
&= k\mathbb{E}\max\left\{\tilde{R}\frac{\lambda}{\sigma} + 1 - \frac{\lambda}{\sigma}, 0\right\} \\
&= \frac{k\lambda}{\sigma}\left(\mathbb{E}\left[\tilde{R}\middle|\tilde{R} > \frac{\lambda-\sigma}{\lambda}\right] - \frac{\lambda-\sigma}{\lambda}\right)\left(1 - F\left(\frac{\lambda-\sigma}{\lambda}\right)\right) \\
&= k\Pi\left(1, \frac{\lambda}{\lambda-\sigma}\right)
\end{aligned}
\tag{10}
$$

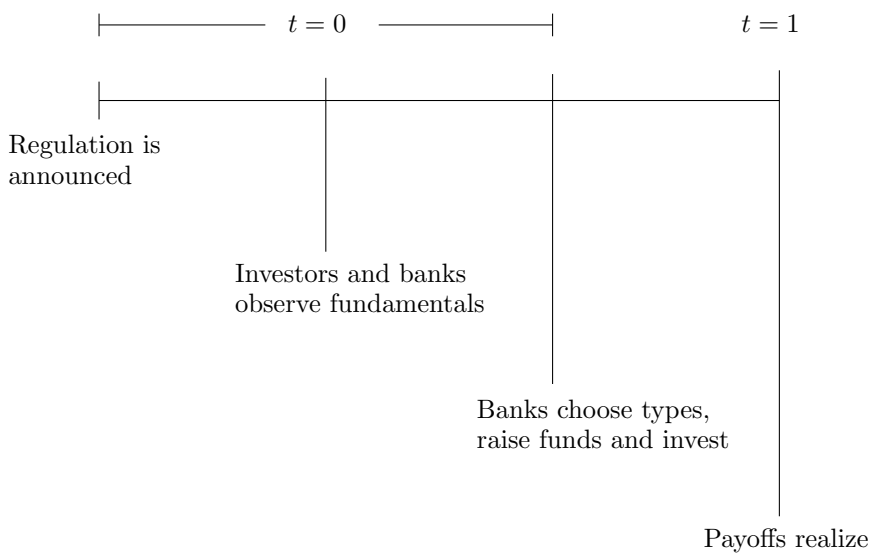
Note that the profit of a commercial bank does not depend on the amount of debt it raises. As discussed earlier, the regulation in this setting amounts to a restriction on how much the bank can invest in the risky project (it is $\alpha(k+d) = \lambda k/\sigma$). Hence, given the capital requirement, the (risk) leverage (λ/σ) of a commercial bank carries no additional information about the riskiness of the bank, for the bank invests all the funds in excess of the limit imposed by the regulation ($k + d - \lambda k/\sigma$) into the riskless project.

A bank decides whether to be SB or CB by comparing the expected values of profits in the two scenarios – $\Pi^{SB}(k)$ and $\Pi^{CB}(k)$. Both these expected profits are linear functions of k (with a positive slope). Therefore, the bank's decision is straightforward:

$$
\begin{aligned}
&\text{If } \Pi\left(1, \tfrac{\lambda}{\lambda-\sigma}\right) > \Pi(r^*, s^*), \quad \text{CB} \\
&\text{If } \Pi\left(1, \tfrac{\lambda}{\lambda-\sigma}\right) < \Pi(r^*, s^*), \quad \text{SB} \\
&\text{If } \Pi\left(1, \tfrac{\lambda}{\lambda-\sigma}\right) = \Pi(r^*, s^*), \quad \text{indifferent}
\end{aligned}
$$

If these profit functions were not linear in bank's capital, the banks will split (if profit is concave in k) or merge (if profit is convex in k) their capital resulting in a degenerate distribution of bank capital. Moreover, to have non-zero mass of both types of banks in equilibrium, these linear functions should coincide – the indifference condition. One important implication of the indifference is that the equilibrium level of risk is indeterminate. As we present in Section 6, the optimal regulation policy resolves this issue by associating the level of capital requirement that supports this indifference with the desired level of risk exposure.

5 Timeline



At $t = 0$, the following events take place in order:

1. The regulator announces the capital requirement schedule $\lambda(\cdot)$.
2. The agents (investors and banks) observe the fundamentals of the economy μ and σ.
3. An equilibrium is reached: (i) The banks choose their types – SB or CB, no bank does better by changing its type, (ii) The SBs raise debt at the interest rate r^* such that their asset to liability ratio is s^*; the SBs invest all their funds in the risky project, (iii) The CBs raise debt at the risk-free interest rate and their risk leverage is λ/σ.

At $t = 1$, all uncertainties are resolved and cashflows are realized.

At any interim $t \in (0, 1)$, if μ or σ changes, the agents repeat the steps 2 and 3 listed above. Since the projects pay off only at $t = 1$ and there is no time value of money, all decisions are made in an identical way irrespective of their timing. In other words, if $(\mu_t, \sigma_t) = (\mu_{t'}, \sigma_{t'})$, the economy looks exactly the same for t and t'. Despite this static nature, we allow the economic fundamentals to change and the agents to immediately respond to the changes in the interim because one of our objectives is to come up with a regulation scheme that achieves its goal (the

optimal aggregate risk exposure) for any realized path of μ_t and σ_t. This setup enables us to avoid bringing in another layer of complexity associated with dynamic programming while it clearly exhibits how our mechanism delivers the robustness.[17] The projects can be invested and disinvested without incurring any loss.

6 Optimal Regulation

We arrive at the optimal regulation policy in three steps: First, we determine the level of capital requirement that makes the banks indifferent between being SB and CB in equilibrium. The important result here is that the capital requirement needs to be relaxed in good times to prevent the banks running into shadows: procyclicality. Second, we relate the relative size of the two banking sectors (the macro variable observed by the regulator) to the aggregate investment in the risky project. The optimal shadow banking sector size is also found to be procyclical. Third, we tie the conditions obtained in the previous two steps to obtain the optimal capital requirement as a schedule of the relative banking sector size. In particular, the regulation is not be specified in terms of the economic fundamentals (μ, σ).

In the following explanation of the mechanism of the optimal regulation, we will set $\sigma = 1$ and let μ vary over time.[18]

[17]There is a caveat in this setup: Since the projects pay nothing interim and the agents can change their decision freely any time, it is not possible to force the agents to behave optimally before $t = 1$. They lose nothing from deferring making a decision in an arbitrary way. For instance, we can completely ignore $t \in [0, \frac{1}{2})$ with no influence on the economy for all $t \in [\frac{1}{2}, 1]$. Note that we adopt this setting solely for the purpose of showing how the regulation is robust to economic fluctuations and what forces take the economy to a new equilibrium. Thus, instead of changing our setting so that procrastination is strictly dominated, we make a behavioral assumption that at any t all the agents participate; that is, they choose optimal actions at any t as if the economy does not fluctuate AND there is no more chance to alter their decisions from time t on. We argue that, even though this assumption is not necessarily true in our model, it is a realistic description of economic agents' behavior most of the time. We are just silent about why it takes place in the market.

[18]We expect that the similar intuition should carry over when σ is varied and μ is held fixed. The example in the next section verifies this intuition. However, the analogous results for this case have been harder to prove analytically.

6.1 Capital Requirement for Banks' Indifference

In order to achieve an equilibrium in which both types of banking sectors have a non-zero mass, a necessary condition is

$$\Pi\left(1, \frac{\lambda}{\lambda - 1}; \mu, 1\right) = \Pi(r^*, s^*; \mu, 1). \tag{11}$$

This constraint gives a mapping

$$\lambda = \Lambda(\mu). \tag{12}$$

If the condition (12) is satisfied, each bank is indifferent between the two types. If $\lambda > \Lambda(\mu)$, all the banks prefer to be CB, and vice versa.

We make a technical assumption on the risky project return distribution F.

Assumption 2. *For all $\hat{R} > 0$, F satisfies*

$$\frac{1 - F(\hat{R})}{f(\hat{R})} > \frac{\int_{\hat{R}}^{\infty} (1 - F(R))\, dR}{1 - F(\hat{R})}. \tag{13}$$

This assumption ensures that the banks' iso-profit curves become flatter as μ increases.

Lemma 7. *Under Assumption 2, MRS^{Π} is decreasing in μ.*

Not all distributions satisfy Assumption 2 but the next lemma states one sufficient condition.

Lemma 8. *If the hazard rate of F is increasing, F satisfies the condition (13).*

Proof. To accommodate the standard notation, define

$$f(\cdot|0) = \frac{1}{\mu}(1 - F(\cdot))$$
$$f(\cdot|1) = f(\cdot)$$

Note that $f(\cdot|0)$ is a density. If the hazard rate is increasing, for all $R_1 > R_0$,

$$\frac{f(R_1|1)}{f(R_1|0)} = \frac{\mu f(R_1)}{1 - F(R_1)} > \frac{\mu f(R_0)}{1 - F(R_0)} = \frac{f(R_0|1)}{f(R_0|0)}.$$

Therefore, $f(R|\theta)$, $\theta = 0, 1$, has the monotone likelihood ratio property (MLRP). As is well known, a necessary condition is that the hazard rate is decreasing in θ: for all R,

$$\frac{f(R|1)}{1 - F(R|1)} < \frac{f(R|0)}{1 - F(R|0)}$$

Substituting the original densities back, we obtain

$$\frac{f(R)}{1 - F(R)} < \frac{1 - F(R)}{\mu - \int_0^R (1 - F(x))dx} = \frac{1 - F(R)}{\int_R^\infty (1 - F(x))dx},$$

leading to the condition (13). □

This condition is standard in mechanism design literature, in which it is a sufficient condition for the increasing virtual valuation.

We are now ready to state one of the main results of our paper.

Proposition 2. *The equilibrium capital requirement is procyclical:* $\Lambda'(\mu) > 0$.

This result makes intuitive sense: As the return prospect of the risky project improves when the banking sector is in equilibrium, *ceteris paribus* it is more profitable for any given bank to be a SB rather than a CB since being a SB provides the bank unrestricted access to the risky project. So, in order to restore the indifference between the two bank types, the CB sector needs

to be made more profitable. It is achieved by loosening the capital requirement constraint, that is, by increasing λ.

Figure 6 illustrates the procyclicality of equilibrium capital requirement. As the mean of the risky return μ goes up, the iso-profit curve of the banks Π becomes flatter and the IR curve s^u of the debt holders shifts downward. Then, it is easily seen from the figure that we expect the intercept of the iso-profit curve to go down, implying the equilibrium λ should go up (the ratio $\lambda/(\lambda-1)$ is decreasing in λ).

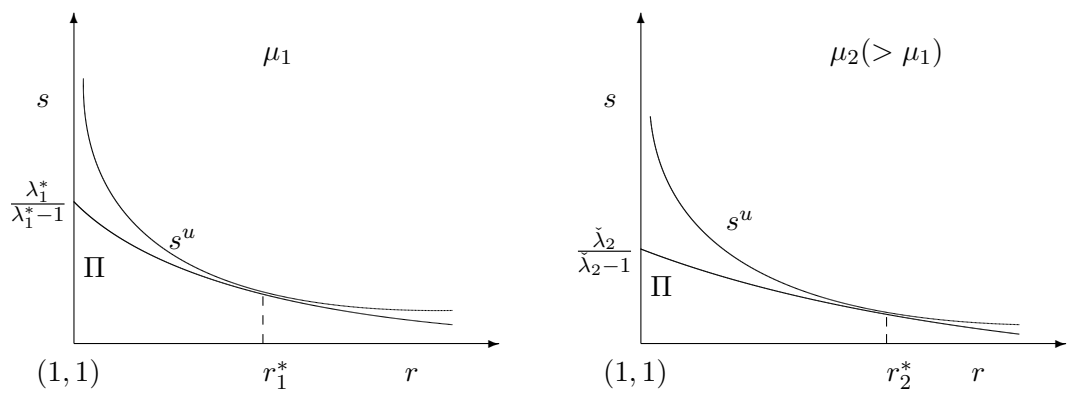


Figure 6: Procyclical Capital Requirement

We have the following comparative statics results

Proposition 3. *When the banks' indifference condition (12) holds, the following variables are increasing in μ.*

1. *Interest rate offered by the shadow banks and their leverage.*
2. *Expected profit of each bank.*
3. *Default probabilities of the shadow banks and the commercial banks.*
4. *Expected cost of deposit insurance per unit of the commercial bank capital.*

The first two results are intuitive. As the investment opportunity improves, a SB is willing to take more leverage by offering a higher interest rate to creditors. With a flatter iso-profit curve, the bank can afford to pay a higher interest rate in exchange for a given increase in the leverage

without hurting its profit. This improvement on the trade-off between raising more debt and paying more interest leads to a procyclical fluctuation of the leverage and interest rate. Also, it is natural that the SBs earn more profit in booms. Since the SBs and the CBs have the same expected profit in equilibrium, the CBs' profit also increases with higher μ.

To see the third result, note that there are two competing forces that affect the SBs' default probability. One is the direct effect of higher μ. With higher μ, the risky project return distribution shifts to the right and the SBs' default probability decreases under a given debt contract. However, there is a force in the opposite direction that is related to the first result. When the SBs pay higher interest rate and are highly levered, they default more often. We find that the second effect dominates the first effect when all the benefit associated with a higher μ accrues exclusively to the SBs and not their debt holders. This leads to a higher default probability of the SBs when μ is higher. However, surprisingly, the same result also holds for the CBs since their risk leverage increases fast enough with μ to dominate the direct effect of the distribution shifting to the right. In this frictionless economy, the force of regulation arbitrage in equilibrium is so strong that even the CBs become riskier – higher likelihood of default – in good times.

This result is a good example that shows how the interaction of the two competing banking sectors leads to a counterintuitive evolution of the banking industry over the business cycle. It is not surprising that the chance of default of the SB sector increases as the return prospect improves. The interesting part is that the government has to allow the CBs to take more risk in order for the CB sector to stay as attractive as the SB sector. Otherwise, the CB sector would fail to survive and the SBs would dominate the banking sector, a situation the regulator finds undesirable. It should be noted that the CB sector is still safer than the SB sector.

The higher default probability of the CB sector has an interesting implication for the expected cost of deposit insurance borne by the regulator. The last result tells us that the relationship takes a specific form: the cyclicality of the expected cost of deposit insurance corresponds to the cyclicality of the default probability. Suppose that there is a constraint on the expected cost of deposit insurance.[19] Since the total expected cost of deposit insurance is proportional to the size of CB sector and the expected cost per unit CB sector capital, in booms the regulator affords to

[19]For example, the congress may put a limit on the expected cost in the government budget.

allow a bigger sector size of SB, which is the condition for policy stability.

Recall that all the above results are derived purely from the banks' indifference condition. However, the indifference condition per se does not pin down the relative size of the two banking sectors and consequently, the aggregate risk. The regulator controls the aggregate risk by adjusting the proportion of the two types of the banks. We now turn to this issue.

6.2 Optimal Relative Banking Sector Size

For this section, we find it more convenient to work with the shadow banks' leverage (assets/equity) $l = \frac{s}{s-1}$. First note that under the banks' indifference condition, the shadow banks are more levered than the commercial banks on risk-adjusted basis:

Lemma 9. $l(\mu) > \Lambda(\mu)$

Proof. Fix μ. Condition (7) implies $r > 1$. Then since $\Pi_r < 0$ and $\Pi_s < 0$, the banks' indifference condition (11) yields $\frac{\lambda}{\lambda-1} > s$ which implies $\lambda < \frac{s}{s-1} = l$. The banks' indifference requires $\lambda = \Lambda$.

□

Corollary 1. *Default probability of the shadow banks is higher than that of the commercial banks.*

Proof. Equation (9) implies that the SBs default when $\tilde{R} < r/s$ and equation (10) implies that the CBs default when $\tilde{R} < (\lambda - 1)/\lambda$. For a given μ, since $r > 1$ and $s < \Lambda/(\Lambda - 1)$, we have $r/s > (\Lambda - 1)/\Lambda$. Since the banks' indifference requires $\lambda = \Lambda$, the inequality implies that the SBs have a bigger region of default.

□

As discussed earlier, maximizing the investors' welfare is equivalent to ensuring the right fraction of the aggregate wealth $W + K$ flows into the risky project. The above lemma implies that for the same amount of equity capital, a shadow bank invests more in the risky project than a commercial bank. Therefore, the aggregate investment in the risky project is one-to-one related to the proportion of the two bank types in the economy. So, the regulator's objective reduces to obtaining the right relative sector size.

We define a state variable, the relative sector size $\xi \equiv \int_{SB} k_j dj / \int_{CB} k_j dj$. We assume that the regulator is able to estimate the capital size of the SB sector $\int_{SB} k_j dj$. Under this assumption, ξ is an observable to the regulator. In our model, the relative net worth of the banking sector, $K/(W+K)$, stays constant at its $t=0$ value $\mathfrak{K}$.[20] The aggregate investment in the risky project is

$$l \int_{SB} k_j dj + \lambda \int_{CB} k_j dj = \frac{K}{\xi + 1} [l\xi + \lambda]$$

Then, the fraction of the aggregate wealth $W+K$ invested in the risky project in equilibrium is

$$\alpha = \frac{\mathfrak{K}(l\xi + \lambda)}{\xi + 1} \tag{14}$$

(Note: $\alpha \in [\mathfrak{K}\lambda, \mathfrak{K}l]$, since $\xi \in [0, \infty]$). The optimal regulation equates the equilibrium fraction with the target fraction $\phi(\mu, 1)$. In the following, we suppress the constant second argument of ϕ for brevity. By setting $\alpha = \phi(\mu)$, we get

$$\xi(\mu) = \frac{\phi(\mu) - \mathfrak{K}\Lambda(\mu)}{\mathfrak{K}l(\mu) - \phi(\mu)} \tag{15}$$

This equation quantifies the target relative sector size. However, there is a caveat – we have to make sure that the right-hand side is finite and non-negative. This is ensured when the model parameters $\mathfrak{K}$ and μ satisfy the restriction:

$$\phi(\mu) \in (\mathfrak{K}\Lambda(\mu), \mathfrak{K}l(\mu)) \tag{16}$$

This restriction is sensible. $\phi(\mu) < \mathfrak{K}\Lambda(\mu)$ means that even when all the banks are CB (realized ξ is zero), the realized investment in the risky project is too high. Note, the regulator cannot set λ lower than $\Lambda(\mu)$ because in that case all the banks will flee to the SB sector and the realized exposure will jump to $\mathfrak{K}l(\mu)$. In the opposite case of $\phi(\mu) > \mathfrak{K}l(\mu)$, even when all the banks are SB (realized ξ is ∞), the realized investment in the risky project is too low. Note that the mutual contracting between SB and their debt holders limits the maximum risk SB can assume.

[20] In an extension of this model to be discussed later, we let $\mathfrak{K}$ evolve over time.

Next, what is a natural direction of $\phi(\mu)$ in the interval (16) for an increase in μ? The most realistic situation is that the optimal exposure is close to the lower limit $\mathfrak{K}\Lambda(\mu)$ when μ is small and it approaches the upper limit $\mathfrak{K}l(\mu)$ as μ becomes bigger (we illustrate this situation using an example in the next section). To be more precise, the relative distance between $\phi(\mu)$ and $\mathfrak{K}\Lambda(\mu)$

$$\frac{\phi(\mu) - \mathfrak{K}\Lambda(\mu)}{\mathfrak{K}l(\mu) - \mathfrak{K}\Lambda(\mu)}$$

is increasing in μ. This is restated as

$$\phi'(\mu) > \frac{l'(\mu)[\phi(\mu) - \mathfrak{K}\Lambda(\mu)] + \Lambda'(\mu)[\mathfrak{K}l(\mu) - \phi(\mu)]}{l(\mu) - \Lambda(\mu)} \tag{17}$$

The next lemma claims that this condition is guaranteed to hold.

Lemma 10. *The range of values of the model parameters $\mathfrak{K}$ and μ for which the conditions (16) and (17) hold simultaneously is non-empty.*

We use this result in the next subsection.

6.3 Optimal Policy Schedule

We now discuss how the regulator could achieve the target relative sector size $\xi(\mu)$ of (15) even when, in equilibrium, the banks are indifferent between the two types. The idea is to make the level of capital requirement λ directly depend on the realized relative sector size ξ such that whenever ξ is suboptimal ($\xi \neq \xi(\mu)$), λ pushes the banks towards the right sector ($\lambda \neq \Lambda(\mu)$).

This goal can be achieved by rearranging (15) to obtain a mapping $\mu = h(\xi)$ and then combine it with (12) to solve for λ

$$\lambda = \Lambda(h(\xi)) \equiv \Omega(\xi) \tag{18}$$

This is our optimal regulation policy specified in terms of the observable ξ. The schedule Ω is independent of the economic fundamental μ which is necessary since the regulator does not observe μ directly. We now state the second main result of our paper.

Proposition 4. *The regulation policy is well-defined and stable:* $\Omega'(\xi) > 0$.

Proof. Using (15), we obtain the first derivative of $\xi(\mu)$

$$\xi'(\mu) = \frac{\mathfrak{K}(l(\mu) - \Lambda(\mu))}{(\mathfrak{K}l(\mu) - \phi(\mu))^2}\left[\phi'(\mu) - \mathfrak{K}\left(\frac{l'(\mu)\xi(\mu) + \Lambda'(\mu)}{\xi(\mu) + 1}\right)\right]$$

Combining this with (15) and (17), we obtain

$$\xi'(\mu) > 0$$

Therefore, the government wants a bigger SB sector as the risky project has a better prospect.[21] Since this first derivative always has the same sign, $h(\cdot)$ is well-defined and so is $\Omega(\cdot)$. The final step is simple: $\Omega'(\xi) = \Lambda'(h(\xi))h'(\xi) = \frac{\Lambda'(h(\xi))}{\xi'(h(\xi))} > 0$ □

To see how this policy achieves its goal, suppose there are too many SBs in the economy, i.e. $\xi > \xi(\mu)$. Then the formulation (18) and the above proposition imply $\lambda > \Omega(\xi(\mu)) = \Lambda(\mu)$. This makes it profitable for the SBs to switch to the CB sector. Although we do not model the process of switching their types, suppose some SBs have moved to the CB sector. This reduces ξ and, under the above schedule, λ. The SBs will continue to convert to CBs until λ is low enough to make the banks indifferent between the two types. At this point, $\lambda = \Lambda(\mu) = \Omega(\xi)$ which implies $\xi = \xi(\mu)$. The above regulation policy is 'stable' in the sense that if the banking sector is perturbed from the equilibrium, the equilibrium is restored automatically – stable equilibrium.

The above proposition also ensures that the economy transitions smoothly from one equilibrium to another as the fundamentals change. As an illustration, suppose μ goes up to μ'. Then, $\lambda < \Lambda(\mu')$ so CBs find it profitable to convert to SBs implying a higher ξ. In turn, it follows that $\Omega(\xi)$ and λ also increase. It means that λ gets closer to $\Lambda(\mu')$. This feedback process stops when

[21] It is interesting to note that the fraction in the risky project, α, increases with μ even if the SB sector size ξ does not increase. This is because the SB sector leverage is increasing in μ ($l'(\mu) > 0$). $\xi'(\mu) > 0$ means that this increase in α is insufficient to obtain the optimal risk exposure $\phi(\mu)$ and that only when both ξ and l are increasing in μ, the realized risk exposure is optimum.

λ has risen to its new equilibrium value $\lambda' = \Lambda(\mu')$ with new mean and, by construction, we reach a new ξ, $\xi' = h^{-1}(\mu')$. Note the critical role of $\Omega'(\xi) > 0$ in transition to the new equilibrium.

7 An Example

We solve for equilibrium when the investors' preference is CRRA and risky asset return follows binomial distribution[22]:

$$u(x) = \frac{x^{1-\gamma} - 1}{1-\gamma}, \ \gamma \neq 1$$

$$\tilde{R} = \begin{cases} \overline{R} & \text{with probability } p, \\ \underline{R} & \text{with probability } 1-p \end{cases}$$

where $p > 0$, $\overline{R} > 1 > \underline{R} > 0$, $p\overline{R} + (1-p)\underline{R} > 1$.

The solution of the program

$$\max_{\alpha} \ \mathbb{E}u\left(1 + \alpha(\tilde{R} - 1)\right)$$

is given by

$$\alpha^* = \phi = \frac{z-1}{\overline{R} - 1 + z(1-\underline{R})},$$

$$\text{where } z = \left[\frac{p(\overline{R}-1)}{(1-p)(1-\underline{R})}\right]^{\frac{1}{\gamma}}$$

The investors' participation constraint is

$$\mathbb{E}u\left(\min\{r, s\tilde{R}\}\right) = u(1). \tag{19}$$

[22]Although binomial distribution does not satisfy the distributional assumptions we imposed, we obtain the same results. The assumptions that binomial distribution does not satisfy – for instance, continuity – are made for the ease of proving results. We believe that the same results hold for discrete and/or non-differentiable distributions which involve more technicalities without much gain.

We will restrict ourselves to the interesting case (see below):

$$\underline{R} < \frac{r}{s} < \overline{R} \tag{20}$$

Then, (19) yields the relation between s and r

$$s^u(r) = \frac{1}{\underline{R}} \left(\frac{1-p}{1-pr^{1-\gamma}} \right)^{\frac{1}{\gamma-1}} \tag{21}$$

(Note: To keep s^u bounded away from 1 as $r \to \infty$, we require $(1-p)^{\frac{1}{\gamma-1}} > \underline{R}$). Shadow bank's profit function

$$\Pi(r,s) = \frac{1}{s-1} \mathbb{E} \max\{s\tilde{R} - r, 0\}$$

takes the form

$$\Pi(r,s) = p \left(\frac{s\overline{R} - r}{s-1} \right).$$

On the investors' participation constraint curve $s^u(r)$, this profit function is

$$\Pi(r) = \Pi(r, s^u(r))$$

FOC: $\Pi'(r) = 0$ yields the solution contract

$$\left(\frac{r^{\gamma-1} - p}{1-p} \right)^{\gamma} = \left(\frac{r^{\gamma} - p\overline{R}}{\underline{R}(1-p)} \right)^{\gamma-1} \tag{22}$$

$$s = \frac{r^{\gamma} - pr}{r^{\gamma} - p\overline{R}} \tag{23}$$

Substituting s from (21) into the left hand-side of (20), we get $1 < r$ and substituting s from (23) into the right hand-side of (20), we get $r < \overline{R}$. So consistent with the assumption (20), we look for the solution(s) of (22) in the range $1 < r < \overline{R}$.

Lemma 11. *(22) has a unique solution in the range* $1 < r < \overline{R}$.

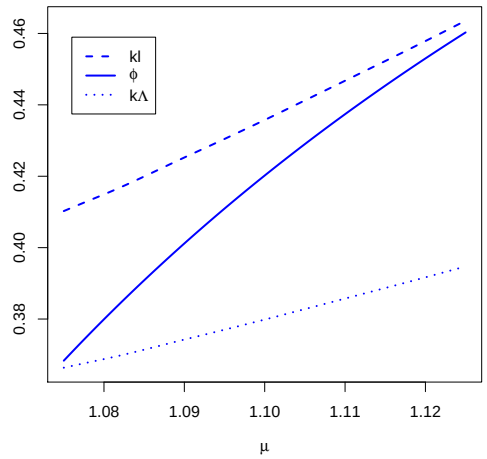


Figure 7: Restriction on μ and $\mathfrak{K}$: conditions (16) and (17), ($\gamma = 4, \sigma = 0.15, \underline{R} = 0.27, \mathfrak{K} = 0.246$)

Equilibrium values of λ and ξ are given by

$$\Pi\left(1, \frac{\lambda}{\lambda - \sigma}\right) = \Pi(r, s) \Rightarrow \frac{\lambda}{\sigma} = \frac{r^\gamma - p}{p(\overline{R} - 1)}$$
$$\xi = \frac{\phi - \mathfrak{K}\lambda/\sigma}{\mathfrak{K}l - \phi} \text{ where } l = \frac{s}{s-1} = \frac{r^\gamma - pr}{p(\overline{R} - r)}$$

Next, we plot the various variables by fixing $\underline{R}$, varying p and $\overline{R}$. We have $\mu = p\overline{R} + (1-p)\underline{R}$ and $\sigma^2 = p(1-p)(\overline{R} - \underline{R})^2$, therefore we write $p(\mu, \sigma) = \frac{(\mu - \underline{R})^2}{(\mu - \underline{R})^2 + \sigma^2}$ and $\overline{R}(\mu, \sigma) = \mu + \frac{\sigma^2}{\mu - \underline{R}}$.

8 Robustness of the policy Ω

The knife-edge nature of the banks' indifference condition, $\lambda = \Lambda(\mu)$, raises the concern that a small error in implementation of the right policy Ω may lead to a catastrophic effect on the economic outcome. In particular, suppose that due to an estimation error in the model parameters (investors' risk aversion coefficient γ, the return distribution F of the risky project etc), the regulator ends up implementing a policy which is *close* to the right policy. In the following, we show that the difference in outcomes under the approximate policy and under the right policy is *bounded*.

Accordingly, suppose there exists a sequence of policies $\{\Omega_n\}$ which is uniformly convergent

to the right policy Ω. Then, for a given $\varepsilon > 0$, we have

$$|\Omega(\xi) - \Omega_n(\xi)| < \varepsilon, \qquad \forall \xi$$

for a sufficiently large n.

At a specific μ, imagine that the banking sector reaches the optimal ξ^* if Ω is implemented. In contrast, if Ω_n satisfying the above condition is implemented, the banking sector will be settled at a suboptimal ξ_n satisfying

$$\Omega(\xi^*) = \Omega_n(\xi_n) = \Lambda(\mu).$$

Note that the banks' indifference condition is $\lambda = \Lambda(\mu)$ irrespective of the regulation policy Ω. If ε is small enough, we can use a first-order approximation:

$$\Omega(\xi_n) = \Omega(\xi^*) + (\xi_n - \xi^*)\Omega'(\xi^*) + o(|\xi_n - \xi^*|)$$

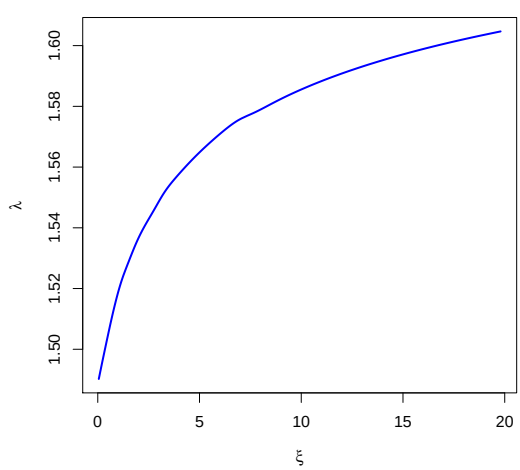


Figure 8: Optimal Policy Ω, $(\gamma = 4, \sigma = 0.15, \underline{R} = 0.27, \mathfrak{K} = 0.246)$

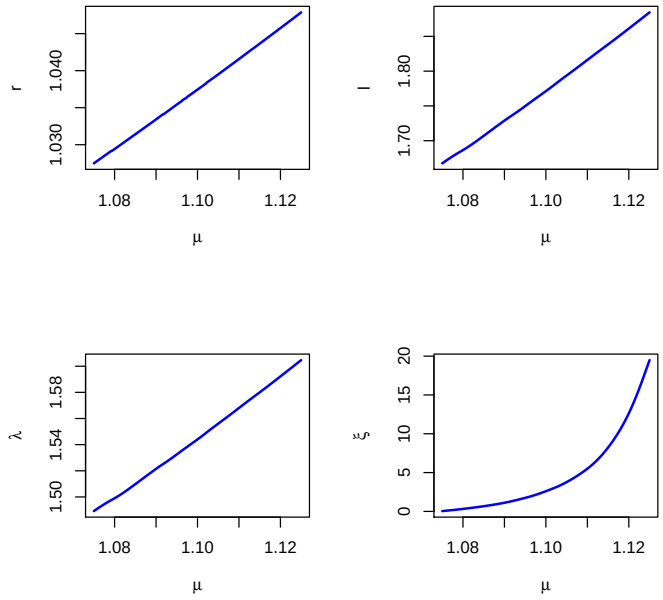


Figure 9: Comparative statics of equilibrium variables r, l, λ, ξ, $(\gamma = 4, \sigma = 0.15, \underline{R} = 0.27, \mathfrak{K} = 0.246)$

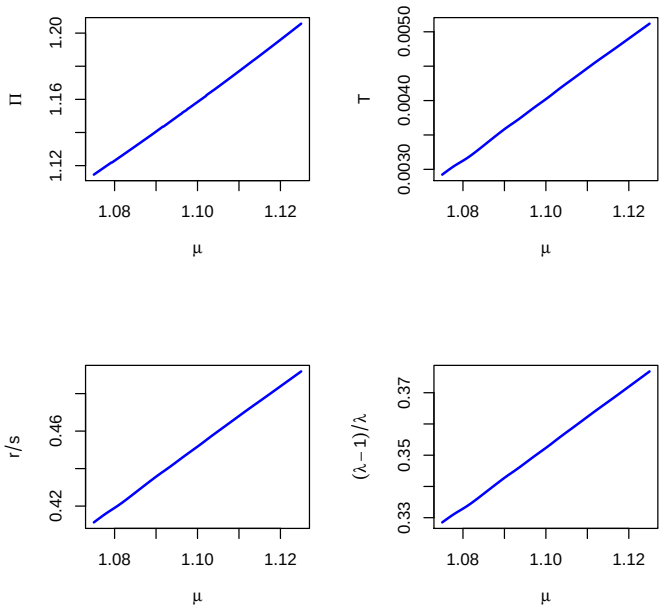


Figure 10: Comparative statics: Banks' Expected Profit Π, Deposit Insurance Payouts T, Default Boundaries of SB: r/s and CB: $(\lambda - 1)/\lambda$, $(\gamma = 4, \sigma = 0.15, \underline{R} = 0.27, \mathfrak{K} = 0.246)$

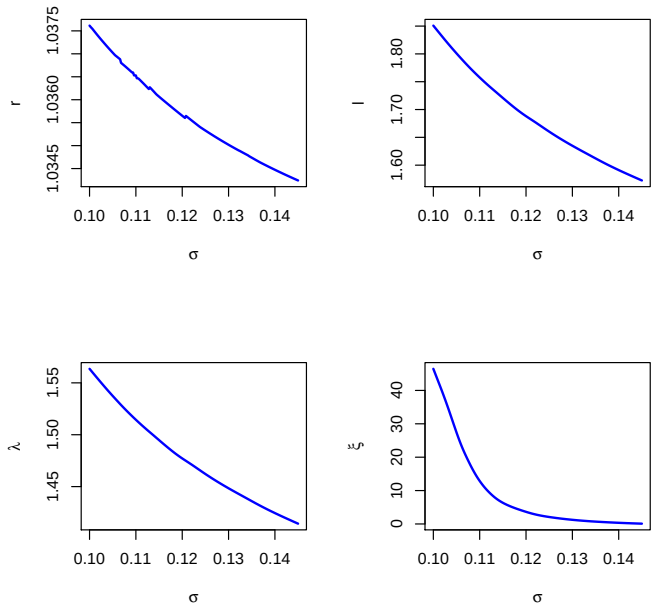


Figure 11: Comparative statics of equilibrium variables r, l, λ, ξ, $(\gamma = 4, \mu = 1.1, \underline{R} = 0.24, \mathfrak{K} = 0.291)$

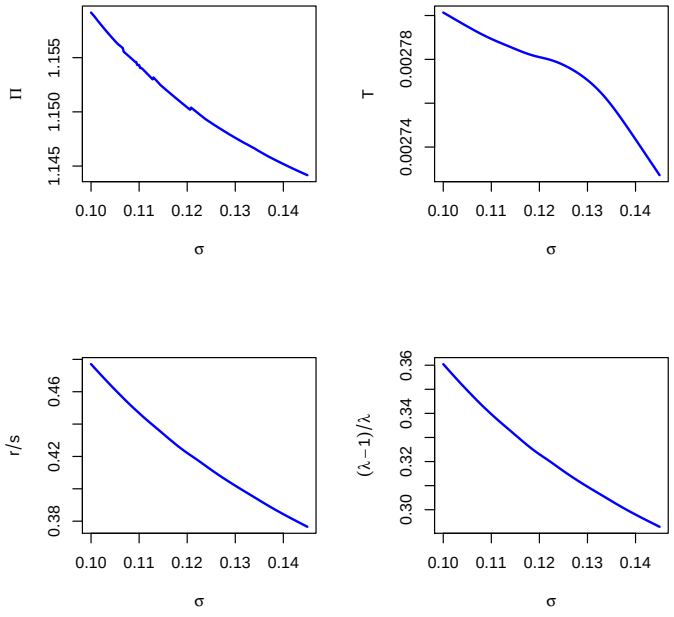


Figure 12: Comparative statics: Banks' Expected Profit Π, Deposit Insurance Payouts T, Default Boundaries of SB: r/s and CB: $(\lambda - 1)/\lambda$, $(\gamma = 4, \mu = 1.1, \underline{R} = 0.24, \mathfrak{K} = 0.291)$

Then,

$$
\begin{aligned}
|\xi_n - \xi^*| &\approx \left| \frac{\Omega(\xi_n) - \Omega(\xi^*)}{\Omega'(\xi^*)} \right| \\
&= \left| \frac{\Omega(\xi_n) - \Omega_n(\xi_n)}{\Omega'(\xi^*)} \right| \\
&< \frac{\varepsilon}{|\Omega'(\xi^*)|}
\end{aligned}
$$

The fraction of the aggregate investment in the risky project ('risk exposure') is

$$
\alpha(\xi) \equiv \frac{\mathfrak{K}(l\xi + \lambda)}{\xi + 1}.
$$

The values of l and λ are invariant for different policies in equilibrium, justifying $\alpha(\cdot)$ is a function of ξ only. The difference in risk exposures under the two policies is approximated using the above results:

$$
\begin{aligned}
|\alpha(\xi_n) - \alpha(\xi^*)| &= \mathfrak{K}(l - \lambda) \frac{|\xi_n - \xi^*|}{(1 + \xi_n)(1 + \xi^*)} \\
&< \mathfrak{K}(l - \lambda) \frac{\varepsilon}{|\Omega'(\xi^*)|}
\end{aligned}
$$

So, if $|\Omega'(\xi^*)|$ is bounded away from zero, the risk exposure is off the target in the order of ε. Ω' is bounded away from zero if $\xi'(\mu)$ is bounded above which, in turn, requires $\mathfrak{K}l - \phi$ is bounded away from zero (see the proof of Proposition 4). This condition comes at almost no cost since $\mathfrak{K}l - \phi > 0$ in the relevant range of μ (see condition (16)).

We can also ask how much loss in the social welfare is incurred from this deviation. The next proposition characterizes the magnitude of this loss.

Proposition 5. *The loss in the social welfare when the approximate policy Ω_n is implemented is an order smaller than the deviation of Ω_n from the right policy.*

Proof. The social welfare is given by

$$
(W + K)\mathbb{E}\left[u\left(1 + \alpha(\tilde{R} - 1)\right)\right].
$$

It is maximized at $\alpha = \alpha(\xi^*)$. The corresponding first order condition is

$$\mathbb{E}\left[u'\left(1+\alpha(\xi^*)(\tilde{R}-1)\right)(\tilde{R}-1)\right] = 0.$$

The welfare loss is

$$\begin{aligned}
&(W+K)\mathbb{E}\left[u\left(1+\alpha(\xi_n)(\tilde{R}-1)\right) - u\left(1+\alpha(\xi^*)(\tilde{R}-1)\right)\right] \\
=&(W+K)\left(\alpha(\xi_n)-\alpha(\xi^*)\right)\mathbb{E}\left[u'\left(1+\alpha(\xi^*)(\tilde{R}-1)\right)(\tilde{R}-1)\right] + O\left((\alpha(\xi_n)-\alpha(\xi^*))^2\right) \\
=&O(\varepsilon^2)
\end{aligned}$$

Therefore, the welfare loss is an order smaller. □

9 Dynamic Environment

In this section, we extend our model to incorporate a mechanism that justifies countercyclical capital requirement policy in the traditional setup. In this modeling environment, we will identify the exact source of the opposite policy suggestion, and the qualifications that are necessary to obtain the countercyclical capital requirement policy.

We consider a stylized dynamic version of our static model with one key difference – the relative net worth of the banking sector, $\mathfrak{K} \equiv K/(W+K)$, evolves over time. In the static setting, we assumed $\mathfrak{K}$ is fixed. However, in a multi-period setting, as the risky project returns realize every period, the wealth of the investors and the banks change, leading to fluctuation in $\mathfrak{K}$. In the following, we delineate how this changes the form of regulation policy.[23]

If $\mathfrak{K}$ changes over time, the optimal ξ depends not only on μ but also on $\mathfrak{K}$ at the time according to (15). This dependence requires us to include $\mathfrak{K}$ as a state variable in our optimal regulation. To keep our analysis simple, we make strong assumptions on how agents behave in the new environment. The key assumption is that the one-period economy repeats every period with the only exception being the agents' wealth levels. Accordingly, given μ, we have the same values for λ, ϕ, r, and l as before since these are independent of $\mathfrak{K}$. We do not prove why this situation

[23]So we ignore the effect of dynamic programming.

is optimal among all possibilities of contracts. As explained, the only part that is affected is the determination of function $h(\cdot)$ in (18): from $h(\xi)$ to $h(\xi, \mathfrak{K})$. To guarantee h is a well-defined function, we need the same condition as Proposition 10:

$$\frac{\partial \xi}{\partial \mu} > 0,$$

because then the Jacobian determinant of $(\xi, \mathfrak{K})'$ is always positive. Although this condition looks the same as in the static case, here it puts a restriction on the realized returns that does not exist in the static setting, because $\mathfrak{K}$ depends on the history of the realized returns and we require $\phi(\mu) \in (\mathfrak{K}\Lambda(\mu), \mathfrak{K}l(\mu))$. In a finite period case, at the minimum, we claim that the above partial derivative is positive if the return distribution has a corresponding upper bound.

The sequence of events in the period t is as follows:

1. The risky project return R_t is realized and observed.
2. $\mathfrak{K}_t$ is determined as a function of R_t and $\mathcal{F}_{t-1}$, where $\mathcal{F}_{t-1}$ is the information set at the end of the period $t-1$.
3. μ_t is observed by the market agents (not the regulator) and the debt contract terms $r(\mu_t)$ and $l(\mu_t)$ are determined.
4. The banks choose their types under the regulation policy $\Omega(\cdot; \mathfrak{K}_t)$.
5. A new equilibrium is achieved. The equilibrium values of λ_t and ξ_t are determined as before.

Before describing the optimal regulation in this dynamic environment, we briefly explain the law of motion of $\mathfrak{K}$ for the sake of completeness. If return R_t is realized at time t (with σ fixed to unity), the return (per unit equity capital) to each bank type is

$$R_t^{SB} \equiv \max\{l_{t-1}R_t - (l_{t-1} - 1)r_{t-1}, 0\}$$
$$R_t^{CB} \equiv \max\{\lambda_{t-1}R_t + 1 - \lambda_{t-1}, 0\}$$

Then the aggregate capital of the banking sector is

$$K_t = K_{t-1}\left[\frac{1}{\xi_{t-1}+1}R_t^{CB} + \frac{\xi_{t-1}}{\xi_{t-1}+1}R_t^{SB}\right].$$

On the other hand, the investors' wealth grows to

$$W_t + K_t = (W_{t-1} + K_{t-1})[1 + \phi_{t-1}(R_t - 1)]$$

because under the optimal capital regulation, ϕ fraction of the aggregate wealth flows into the risky project. Then

$$\begin{aligned}\mathfrak{K}_t &= \frac{K_t}{W_t + K_t} \\ &= \mathfrak{K}_{t-1}\left[\frac{R_t^{CB} + \xi_{t-1}R_t^{SB}}{(\xi_{t-1}+1)[1+\phi_{t-1}(R_t-1)]}\right]\end{aligned}$$

As described above, we now have a new policy form, $\Omega(\xi; \mathfrak{K})$. It should be noted that the roles played by ξ and $\mathfrak{K}$ for this policy are very different. Since $\mathfrak{K}_t$ is determined ex-ante, it plays no role in implementing stable and robust policy. Effectively, we have a menu of the policies $\Omega(\xi)$, one for each value of $\mathfrak{K}_t$ which is known perfectly before the banks make any decision at time t. In contrast, ξ is the vehicle that drives the economy to a fixed point which is optimal by construction.

The optimal capital requirement is procyclical even in this environment. Cyclicality is purely determined by the sign of Λ', which is independent of $\mathfrak{K}$. In our model, the fundamental force driving the cyclicality of capital requirement is the strategic behavior of the banks that they can choose to be regulated or unregulated. Hence, the cyclicality comes prior to considering implementing the optimal risky investment. Admittedly our model has a structure to manifest the force for cyclicality, but it is also true that this force is hard to annihilate.

Kashyap and Stein (2004) propose countercyclical capital requirement policy based on the argument that the shadow value of bank capital rises in recessions. As they point out, it amounts to the claim that the effect of capital crunches on the shadow value dominates the effect of deteriorated investment opportunities on the shadow value. Therefore, they conclude that the capital requirement should be loosened during recessions to alleviate the excessive scarcity of bank capital in bad times.

Our model generates countercyclical capital requirement policy as in Kashyap and Stein (2004) when we shut down the shadow banking sector (impose $\xi = 0$) and let the regulator observe the

business cycle (μ_t) directly. In this case, the regulator achieves the optimal level of investment in the risky project by controlling λ directly (see equation (14)):

$$\lambda_t = \frac{\phi(\mu_t)}{\mathfrak{K}_t}$$

Now suppose that, in the spirit of Kashyap and Stein (2004), in bad times for a fixed λ, the banks' lending $\lambda\mathfrak{K}_t$ shrinks too much to achieve the desirable level of economic activity $\phi(\mu_t)$, although $\phi(\mu_t)$ also goes down. That is, we assume $\phi(\mu_t)$ is more stable than $\mathfrak{K}_t$:[24]

$$\frac{d\mathfrak{K}_t/\mathfrak{K}_t}{d\mu_t/\mu_t} > \frac{d\phi(\mu_t)/\phi(\mu_t)}{d\mu_t/\mu_t} \tag{24}$$

Then, in order to reach the optimal risk exposure, $\phi(\mu_t)$, the regulator needs to increase λ_t – a looser capital requirement in bad times. To see more formally that the optimal capital requirement is countercyclical, note that

$$\begin{aligned}\frac{d\lambda_t}{d\mu_t} &= \frac{1}{\mathfrak{K}_t^2}\left[\mathfrak{K}_t\frac{d\phi(\mu_t)}{d\mu_t} - \phi(\mu_t)\frac{d\mathfrak{K}_t}{d\mu_t}\right] \\ &= \frac{\phi(\mu_t)}{\mu_t\mathfrak{K}_t}\left[\frac{d\phi(\mu_t)/\phi(\mu_t)}{d\mu_t/\mu_t} - \frac{d\mathfrak{K}_t/\mathfrak{K}_t}{d\mu_t/\mu_t}\right] < 0\end{aligned}$$

where the last inequality follows from (24). This countercyclical policy mitigates a credit-freeze in bad times and a credit-craze in good times.

In our model, this advocacy of countercyclical capital requirement can be valid only if there is no alternative type of banks – the unregulated banks. Once we open the conduit for capital flow into the shadow banking sector, the effects in the literature disappear due to the competition between the two types of banks. Instead, with our optimal regulation, we showed $\frac{d\lambda_t}{d\mu_t} > 0$. Although this comparison requires some analogy and interpretation, our model patently shows how the consideration of strategic behavior of banks can result in very different policy prescription.

[24]Note that in this setting the bank capital $\mathfrak{K}_t$ varies with the business cycle μ_t.

10 Crisis

One of the main advantages to implement a market-based regulation is that the economy can more flexibly respond to changes in market conditions. This benefit cannot be obtained by a conventional regulatory framework based on agent(firm)-specific information. It leads us to consider how our new approach performs in case of a crisis.

Even though our model is static, its construction bears the idea of finding an optimal regulation under fluctuating fundamentals. We can imagine that, if the economy is Markovian and other parameters except for the expected return do not change over time, the equilibrium condition from the profit comparison and the optimality condition from the government's objective would lead us to a regulation very similar to what we obtained in previous sections. Instead of solving for the regulation in a dynamic setting, however, we assume that $\Omega(\xi)$ is the optimal regulation given other parameters in a dynamic setting.

Turning to modeling a crisis, we find two most prevailing approaches in the literature: (i) The first path introduces a market friction or constraint on resource allocation (mostly borrowing constraints) into a standard model and investigate what happens to variables of interest such as prices when the constraint binds. (ii) The other approach assumes that agents have heterogeneous beliefs and see how the beliefs lead to bubbles and/or crashes. Since the main focus of this paper is not crisis, however, we limit ourselves to a (very) reduced form modeling of crisis in this paper. We focus on the fact that one of the notable features of crisis is a 'flight to quality'. In our model, the feature can be interpreted that the investors tilt their portfolio towards the commercial banks during crisis even if the economic fundamentals do not move. We incorporate the investors' behavior through time-varying risk aversion (γ).

We illustrate two approaches to hedge against this fluctuating risk aversion: with and without a detection technology of crisis. Neither of the approaches can be a perfect instrument, so we call for different assumptions for each of the approaches to work. With a detection technology, the government is accurately informed of the timings of the inception and the end of a crisis and implements a different regulation during the crisis. The crisis regulation will enable the economy to achieve the target risk exposure for different γ. In contrast, while the second approach does not require the detection technology in place, it assumes that risk aversion always moves slowly

compared to the economic fundamentals and make use of past information.

10.1 With a detection technology of crisis

The first approach to contain a crisis is to think of crisis as a transitory drift in the investors risk aversion parameter γ from its pre-crisis value γ_0.[25] We assume that even when the government is able to exogenously detect this jump, it cannot track γ as it varies during the crisis before stabilizing to the end of crisis value γ_1 (potentially same as γ_0).[26] It is hard to imagine that the government can observe time-varying risk aversion contemporaneously. We further assume that the duration of crisis is short and the fundamental, μ, does not vary much during the crisis. Then, the fluctuation of preference for safety is much starker than that of fundamentals over the period and it would be more beneficial to implement a regulation immune to the changing liquidity demand rather than to the changing fundamentals.

Now suppose the government implements $\Omega_{\gamma_0}(\xi)$ at time 0. The regulation is indexed by the risk aversion coefficient in order to make clear that the optimal regulation function depends on the parameter (at time 0, the coefficient is known as γ_0). At time t_1, the following sequence of event unfolds: (i) ξ_{t_1} is realized and observed, (ii) a crisis occurs (iii) the government implements a new policy. Since the government observes ξ_{t_1} before the crisis happens, it knows μ_{t_1}. The assumption we need to proceed is that $\mu_t = \mu_{t_1}$ until the crisis is over. Certainly this assumption will cause some loss of efficiency in the regulations on and after, but the very form of our regulation, which is the dependence upon ξ, limits the deviation from optimality.

Given that μ is fixed, l and r are functions of only γ_t, implying that (12) becomes

$$\lambda = \Gamma^C(\gamma),$$

for some function Γ^C where the superscript denotes a crisis. In the same manner, (15) is now

[25] There could many other ways to model a non-standard preference for safety but, as long as the preference is parameterized by a single parameter, the same intuition carries over.

[26] We do not explicitly model the detection mechanism of a crisis here. In the current setup, we can think of a statistical detection of a crisis by looking at the time series of realized ξ. If the government observes a sudden jump in ξ, it can suspect a shock to the economy which is not due to change in fundamentals μ, σ. The detection problem could be interesting on its own, but we do not deal with it further and assume that the government has a device in place to detect whether the economy is in a crisis or not with no lag.

written as

$$\phi(\gamma;\mu_{t_1}) = \frac{\mathfrak{K}}{\xi+1}\left[l(\gamma;\mu_{t_1})\xi + \lambda\right]$$

Following the same strategy in Section 6, these two relations provides a regulation

$$\lambda = \Omega^C_{\mu_{t_1}}(\xi)$$

Even though the regulations look the same, their role is quite different. In normal times, $\Omega_\gamma(\xi)$ allows the economy achieve the target exposure for any state of fundamentals. In contrast, in a crisis, $\Omega^C_{\mu_{t_1}}(\xi)$ is implemented to keep the economy at the desirable risk exposure irrespective of the investors' unstable demand for safety.

The crisis rule $\Omega^C_{\mu_{t_1}}$ is implemented at time t_1 until the government is informed that the crisis is over. When it is over at time t_2, by the same logic, the government infers the risk aversion γ_{t_2} at the moment and implements the normal rule $\Omega_{\gamma_{t_2}}(\xi)$.

Certainly μ is also fluctuating during the crisis, and the regulation $\lambda = \Omega^C_{\mu_{t_1}}(\xi)$ does not exactly track the desired level of exposure. Nevertheless, the previous section implies that this deviation is not large as long as the assumption on crisis is valid.

10.2 Without a detection technology of crisis

Here we assume that time is discrete. To be more precise, the frequency of financial reporting of banks is finite. To do without a detection technology, we instead assume that risk aversion fluctuates in a relatively slow manner compared to the economic fundamentals. Although it sounds almost innocuous, this assumption stands in the opposite spirit of the first approach above, in which we hypothesized that μ is more stable than γ in a crisis. It should be noted that γ should be viewed in a broader context, rather than the literal meaning of risk aversion. The parameter summarizes the investors' behavior during a crisis that cannot be explained by fundamentals. Hence, we argue that the two approaches have their own merits and limitations.

Given the slow-moving risk aversion, in order to restrain the effect of changing risk aversion, we can expand our set of state variables being used to ξ_t, ξ_{t-1}, s_{t-1}. It is very costly or even impossible

for the government to observe the contemporaneous ξ_t or s_t, but their past values, ξ_{t-1} and s_{t-1} are readily observable to the extent that the banks report their financial statements truthfully. This past information enables the government to calculate μ_{t-1} and γ_{t-1}. The assumption of stable risk aversion takes the form of $\gamma_t = \gamma_{t-1}$ at time t. Then the optimal policy is to implement the same policy established in Section 6, $\Omega(\xi_t; \gamma_t) = \Omega(\xi_t; \gamma_{t-1})$ in period t. Since γ_{t-1} is a known function of ξ_{t-1} and s_{t-1}, we can equivalently write the optimal regulation as

$$\lambda_t = \Omega(\xi_t, \xi_{t-1}, s_{t-1})$$

It should be noted that the function $\Omega(\cdot)$ does not depend on time, so all the regulator has to do is to announce the schedule $\Omega(\cdot)$ at time 0, as before. In this approach, the deviation comes from the drift of risk aversion over one period, which was assumed to be small.

11 Conclusion

In this paper, we model the decision of banks to become regulated in the form of capital requirement. In making this decision, the banks compare the cost of being regulated – limit on risk exposure – with the benefit – access to cheaper funding. We propose a capital regulation policy that controls the aggregate risk assumed by the banking sector taking into account the fact that the banks may not find it profitable to be regulated. The solution policy obtains the right mix of the risky unregulated (shadow) banking sector and the safe regulated (commercial) banking sector. The optimal policy is shown to be procyclical – relaxed capital requirement during booms and vice versa. This is opposite to the countercyclical policy proposed in literature. The reason for this dichotomy is that the literature ignores the ability of the banks to switch in and out of the regulation purview, and our paper is a first step in this direction.

In our model, the optimal regulation policy achieves the right mix of the two banking sectors even when, in equilibrium, the banks are indifferent between being regulated and unregulated. The idea here is to make the level of capital requirement a schedule of the relative size of the two banking sectors realized at a given time such that whenever the relative size is suboptimal, the level of capital requirement pushes the banks towards the right sector.

The proposed policy has several desirable features: (i) Macroprudential – The objective of the policy is to control systematic risk (rather than focusing on health of each individual bank), (ii) Market-based – The policy instrument is specified as a schedule of a market variable; once the policy schedule is announced, market self-adjusts to obtain the optimal risk exposure, (iii) Robust to business cycle fluctuations – The policy schedule does not depend on the underlying economic fundamentals; under the policy, the economy transitions smoothly from one equilibrium to another as the underlying fundamentals change, and (iv) Robust to measurement errors – Small measurement errors either in risk aversion or risky return distribution on part of the government lead to welfare loss that is an order smaller.

The main motivation behind this paper is the arbitrage of the capital requirement regulation that banks engaged in before and during the crisis of 2007-2009. We believe some banks will always exist in shadows or equivalently, it would be prohibitively costly for the government to employ a loose enough regulation that no bank prefers to be in shadows.[27] Moreover, insuring a huge commercial banking sector incurs a huge direct cost to the government. So, rather than the overarching ambition of regulating everyone (as is the spirit of Dodd-Frank Act 2010), our policy only indirectly influences the size of shadow banking sector to control the systematic risk of the economy.

In a stylized multi-period extension of our model, we introduce one force that obtains the countercyclical capital regulation policy suggested in the literature: dynamic evolution of banking sector capital with business cycle. We show how, after an appropriate adjustment in the implementation of the policy in this framework, the optimal policy is countercyclical if the shadow banking sector is turned-off, while it is procyclical otherwise (that is, when the banks are allowed to be in the shadow). So, our procyclical result holds over and above the traditional argument of countercyclical policy.

[27] We saw that, in Section 6, it is always optimal for the government to have both sectors exist in equilibrium, unless the desired risk exposure is very low.

12 Appendix

Proof of Lemma 2 Define the cut-off return $\hat{R} = \frac{r}{s}$ and

$$EU(r,s) \equiv \int_0^{\hat{R}} u(sR)dF + u(r)(1 - F(\hat{R}))$$

The partial derivatives of EU are

$$\frac{\partial}{\partial s} EU(r,s) = \int_0^{\hat{R}} u'(sR)RdF > 0$$
$$\frac{\partial}{\partial r} EU(r,s) = u'(r)(1 - F(\hat{R})) > 0$$

By definition of $s^u(r)$

$$EU(r, s^u(r)) = 0$$

The marginal rate of substitution (MRS^U) is obtained from the above partial derivatives:

$$MRS^U = -\frac{ds^u(r)}{dr} = \frac{\partial EU/\partial r}{\partial EU/\partial s} = \frac{u'(r)(1 - F(\hat{R}))}{\int_0^{\hat{R}} u'(sR)\, RdF} > 0$$

Turning to convexity, we compute the numerator of $\frac{d^2 s^u(r)}{dr^2}$ and check its sign:

$$\left[u'(r)f(\hat{R})\hat{R}' - u''(r)(1 - F(\hat{R}))\right] \int_0^{\hat{R}} u'(sR)\, RdF$$
$$+u'(r)(1 - F(\hat{R}))\left[u'(r)\hat{R}f(\hat{R})\hat{R}' + s' \int_0^{\hat{R}} u''(sR)\, R^2 dF\right],$$

where $x' \equiv \frac{dx}{dr}$. Since $\hat{R}' = \frac{1}{s} - \frac{r}{s^2}s' > 0$, each of the above terms is positive.

Proof of Lemma 3 In the proof of Lemma 2, we saw

$$\frac{\partial}{\partial s} EU(r,s) > 0$$
$$\frac{\partial}{\partial r} EU(r,s) > 0$$

Therefore,

$$EU(r,s) \geq \lim_{s \to 1} EU(r,s) = \int_0^r u(R)dF + u(r)(1 - F(r)).$$

The limit is denoted by $EU(1,r)$.
Suppose $\lim_{r\to\infty} EU(r,1) = \int_0^\infty u(R)dF > 0$. By continuity, there exists $\overline{r} \in (1,\infty)$, such

that

$$EU(\overline{r},1) = \int_0^{\overline{r}} u(R)dF + u(\overline{r})(1-F(\overline{r})) \geq 0$$

Since $\frac{\partial EU}{\partial s} > 0$,

$$\forall s, \quad EU(\overline{r},s) > EU(\overline{r},1) \geq 0$$

This implies that, for all r such that $EU(r,1) \geq 0$, there is no solution for s that satisfies the investors' IR. In other words, r is upper bounded.
Next, suppose $\lim_{r\to\infty} EU(r,1) = \int_0^\infty u(R)dF < 0$. Then, there exists $\underline{s}$ such that

$$\lim_{r\to\infty} EU(r,\underline{s}) = \int_0^\infty u(\underline{s}R)dF = 0$$

It follows that, for all $s \leq \underline{s}$,

$$EU(r,s) < \lim_{r\to\infty} EU(r,s) = \int_0^\infty u(sR)dF \leq 0.$$

Hence, given s no bigger than $\underline{s}$, the investors' IR condition cannot be satisfied for any r, leading to the conclusion that s is bounded away from 1.

Proof of Lemma 4

$$\Pi(r,s) = \frac{s}{s-1}\left[\int_{\hat{R}}^\infty (R-\hat{R})dF\right],$$

where $\hat{R} \equiv \frac{r}{s} < r$.
We impose a participation constraint

$$\Pi(r,s) \geq \mu \qquad \forall s, r \tag{25}$$

It means that no bank borrows from the investors unless it yields more expected profit than autarky.
A useful implication is

$$\mu \geq r\int_{\hat{R}}^\infty dF \tag{26}$$

First note that $\int_{\hat{R}}^\infty RdF$ is decreasing in $\hat{R}$. Therefore,

$$\mu = \int_0^\infty RdF \geq \int_{\hat{R}}^\infty RdF$$

It follows that

$$\begin{aligned}
\mu &\leq \frac{s}{s-1}\left[\int_{\hat{R}}^{\infty}(R-\hat{R})dF\right] \qquad \text{by (25)} \\
&= \frac{s}{s-1}\int_{\hat{R}}^{\infty} RdF - \frac{s}{s-1}\hat{R}\int_{\hat{R}}^{\infty} dF \\
&\leq \frac{s\mu}{s-1} - \frac{r}{s-1}\int_{\hat{R}}^{\infty} dF,
\end{aligned}$$

leading to the inequality.

To compute the marginal rate of substitution, we first obtain the partial derivatives of $\Pi(r,s)$. The derivatives establish that $\Pi(r,s)$ is decreasing in s and r.

$$\begin{aligned}
\frac{\partial \Pi(r,s)}{\partial s} &= -\frac{1}{(s-1)^2}\left[\int_{\hat{R}}^{\infty}(R-\hat{R})dF\right] + \frac{r}{s(s-1)}\int_{\hat{R}}^{\infty} dF \\
&= -\frac{1}{s(s-1)}\Pi(r,s) + \frac{r}{s(s-1)}\int_{\hat{R}}^{\infty} dF \\
&\leq -\frac{\mu}{s(s-1)} + \frac{r}{s(s-1)}\int_{\hat{R}}^{\infty} dF \qquad \text{by (25)} \\
&\leq 0 \qquad \text{by (26)}
\end{aligned}$$

Also, note that, from the above derivation, we can write

$$\frac{\partial \Pi(r,s)}{\partial s} = -\frac{1}{(s-1)^2}\int_{\hat{R}}^{\infty}(R-r)dF,$$

which is proven to be negative, implying that

$$\int_{\hat{R}}^{\infty}(R-r)dF > 0.$$

Turning to r,

$$\frac{\partial \Pi(r,s)}{\partial r} = -\frac{1}{(s-1)}\left[\int_{\hat{R}}^{\infty} f(R)dR\right] \leq 0$$

It follows that the marginal rate of substitution (MRS^{Π}) is given by

$$MRS^{\Pi} = (s-1)\frac{1-F(\hat{R})}{\int_{\hat{R}}^{\infty}(R-r)dF}, \tag{27}$$

and positive since

$$\int_{\hat{R}}^{\infty}(R-r)dF > 0.$$

as shown above.

The decreasing MRS^{Π} follows from

$$\frac{d}{dr}MRS^{\Pi} = -(s-1)\frac{f(\hat{R})\hat{R}'\int_{\hat{R}}^{\infty}\left(R-\hat{R}\right)dF}{\left[\int_{\hat{R}}^{\infty}(R-r)dF\right]^2} < 0$$

Proof of Lemma 5 Rewrite condition 1,

$$\begin{aligned}
-\frac{ds^u(r)}{dr} &= \frac{\partial\Pi(r,s^u(r))/\partial r}{\partial\Pi(r,s^u(r))/\partial s} \\
&= (s^u(r)-1)\frac{1-F(\hat{R})}{\int_{\hat{R}}^{\infty}(R-r)dF}
\end{aligned}$$

The left hand side is MRS^U and right hand side is MRS^{Π} with $s = s^u(r)$. So, conditions 1 and 2 are equivalent.

Setting $MRS^U = MRS^{\Pi}$, gives us

$$\frac{u'(r)}{\int_0^{\hat{R}} u'(sR)\,RdF} = \frac{s-1}{\int_{\hat{R}}^{\infty}(R-r)dF}$$

We know

$$\begin{aligned}
\Pi(r,s) &= \frac{s}{s-1}\int_{\hat{R}}^{\infty}(R-\hat{R})dF \\
&= \frac{s}{s-1}\int_{\hat{R}}^{\infty}(R-r)dF + \frac{s}{s-1}(r-\hat{R})(1-F(\hat{R})) \\
&= \frac{s}{s-1}\int_{\hat{R}}^{\infty}(R-r)dF + r(1-F(\hat{R}))
\end{aligned}$$

When $MRS^U = MRS^{\Pi}$, $\Pi(r,s)$ can be written as

$$\begin{aligned}
\Pi(r,s) &= s\int_0^{\hat{R}}\frac{u'(sR)\,R}{u'(r)}dF + r(1-F(\hat{R})) \\
&= r\int_0^{\hat{R}}\left(\frac{u'(sR)\,sR}{u'(r)r}-1\right)dF + r \\
&= r\int_0^{r}\left(\frac{u'(x)x}{u'(r)r}-1\right)f\left(\frac{x}{s}\right)\frac{1}{s}dx + r
\end{aligned}$$

Now set $s = s^u(r)$ to obtain the equivalence of conditions 2 and 3

$$MRS^U = MRS^{\Pi}\text{and } s = s(r) \Longleftrightarrow \Pi(r) = J(r)$$

Proof of Lemma 6

$$J(r) = r\int_0^{\hat{R}} \left(\frac{u'(s^u R)s^u R}{u'(r)r} - 1\right) dF + r$$

$$\begin{aligned}
J'(r) &= \int_0^{\hat{R}} \left(\frac{u'(s^u R)s^u R}{u'(r)r} - 1\right) dF - r\int_0^{\hat{R}} \frac{u'(s^u R)s^u R}{[u'(r)r]^2}[u''(r)r + u'(r)]dF \\
&+ r\int_0^{\hat{R}} \frac{s^{u\prime}}{u'(r)r}[u''(s^u R)s^u R^2 + u'(s^u R)R]dF + 1 \\
&= \int_0^{\hat{R}} \left(\frac{u'(s^u R)s^u R}{u'(r)r} - 1\right) dF - r\int_0^{\hat{R}} \frac{u'(s^u R)s^u R}{[u'(r)r]^2}[u''(r)r + u'(r)]dF \\
&- (1 - F(\hat{R}))\frac{\int_0^{\hat{R}}[u''(s^u R)s^u R^2 + u'(s^u R)R]dF}{\int_0^{\hat{R}} u'(s^u R)RdF} + 1 \\
&= \int_0^{\hat{R}} \frac{u'(s^u R)s^u R}{u'(r)r}dF - F(\hat{R}) - r\int_0^{\hat{R}} \frac{u'(s^u R)s^u R}{[u'(r)r]^2}[u''(r)r + u'(r)]dF \\
&- (1 - F(\hat{R}))\frac{\int_0^{\hat{R}} u''(s^u R)s^u R^2 dF}{\int_0^{\hat{R}} u'(s^u R)RdF} - (1 - F(\hat{R})) + 1 \\
&= \int_0^{\hat{R}} u'(s^u R)s^u RdF\left[\frac{1}{u'(r)r} - \frac{u''(r)}{u'(r)} - \frac{1}{u'(r)r}\right] - (1 - F(\hat{R}))\frac{\int_0^{\hat{R}} u''(s^u R)s^u R^2 dF}{\int_0^{\hat{R}} u'(s^u R)RdF} \\
&= -\frac{u''(r)}{u'(r)}\int_0^{\hat{R}} u'(s^u R)s^u RdF - (1 - F(\hat{R}))\frac{\int_0^{\hat{R}} u''(s^u R)s^u R^2 dF}{\int_0^{\hat{R}} u'(s^u R)RdF}
\end{aligned}$$

Both terms are positive, so $J' > 0$ for all r. Other two properties are easily verified.

Proof of Proposition 1 We establish the limits of $\Pi(r)$ at each end. Consistent with our intuition,

$$\lim_{r\to 1} \Pi(r) = \int_0^{\infty} RdF = \mu$$

We use Proposition 3 to find the other limit. Since s is bounded away from 1, $\frac{s}{s-1}$ is upper bounded. Therefore,

$$\lim_{r\to\infty} \Pi(r) = 0$$

Since $H(r) \equiv \Pi(r) - J(r)$ is a continuous function, with $\lim_{r\to 1} H(r) = \mu - 1 > 0$, and $\lim_{r\to\infty} H(r) = -\infty$, there exists a finite r^* that satisfies $H(r^*) = 0$.
To see uniqueness, note that $H'(r^*) = -J'(r^*) < 0$. So, H cuts the x-axis once from above.

It still remains to be shown that the unique extremum is a maximum, not a minimum. If it is a minimum, for big r, $\Pi(r)$ is increasing in r. Since $\lim_{r\to\infty} \Pi(r) = 0$, it follows that

$\Pi(r) < 0$ for some r, an impossibility. So $\Pi(r)$ has a unique solution for $\Pi'(r) = 0$, which is a maximum.

Proof of Lemma 7 We rewrite (27) using a conditional expectation:

$$\begin{aligned} MRS^{\Pi} &= -\frac{s-1}{\frac{\int_{\hat{R}}^{\infty} RdF}{\int_{\hat{R}}^{\infty} dF} - r} \\ &= -\frac{s-1}{\mathbb{E}[\tilde{R}|\tilde{R} \geq \hat{R}] - r} \end{aligned} \tag{28}$$

To see the effect of changing μ, we use the parameterized distribution, f^{μ}, as defined in the setup. Changing μ means horizontal translation, so $f^{\mu} = f(x - \mu)$. That is, all the distributions belong to the location family with f. We rewrite (28) with μ explicitly:

$$MRS^{\Pi} = -\frac{s-1}{\mathbb{E}^{\mu}[\tilde{R}|\tilde{R} \geq \hat{R}] - r}$$

If the conditional expectation $\mathrm{E}^{\mu}[R|R \geq \hat{R}]$ is increasing in μ, the iso-profit curve becomes flatter as μ goes up. The sign of $\frac{\partial \mathrm{E}[R|R \geq \hat{R}]}{\partial \mu}$ is determined by the numerator of the derivative:

$$\begin{aligned} &-\int_{\hat{R}}^{\infty} Rf'(R-\mu)dR \cdot \int_{\hat{R}}^{\infty} f(R-\mu)dR + \int_{\hat{R}}^{\infty} Rf(R-\mu)dR \cdot \int_{\hat{R}}^{\infty} f'(R-\mu)dR \\ &= \hat{R}f(\hat{R}-\mu)\int_{\hat{R}}^{\infty} f(R-\mu)dR + \left[\int_{\hat{R}}^{\infty} f(R-\mu)dR\right]^2 - f(\hat{R}-\mu)\int_{\hat{R}}^{\infty} Rf(R-\mu)dR \\ &= \left[\int_{\hat{R}}^{\infty} f^{\mu}(R)dR\right]^2 - f^{\mu}(\hat{R})\int_{\hat{R}}^{\infty} (R-\hat{R})f^{\mu}(R)dR \\ &= \left[1 - F^{\mu}(\hat{R})\right]^2 + f^{\mu}(\hat{R})(R-\hat{R})\left(1 - F^{\mu}(R)\right)\Big|_{\hat{R}}^{\infty} - f^{\mu}(\hat{R})\int_{\hat{R}}^{\infty} \left(1 - F^{\mu}(R)\right) dR \\ &= \left[1 - F^{\mu}(\hat{R})\right]^2 - f^{\mu}(\hat{R})\int_{\hat{R}}^{\infty} \left(1 - F^{\mu}(R)\right) dR, \end{aligned} \tag{29}$$

where the first and third equalities come from the integration by parts. If (29) is positive, the iso-profit curve becomes flatter for higher μ.

Proof of Proposition 2 We have to consider two effects: changing MRS^{Π} and shifting the investors' IR. First, consider the effect on MRS^{Π} holding IR fixed. Lemma 7 shows that, if the condition (13) holds, the iso-profit curve flattens. Then, $\Pi(r, s; \mu)$ satisfies the *strict Spence-Mirrlees condition* in Edlin and Shannon (1998b). Edlin and Shannon (1998a) show that the strict Spence-Mirrlees condition implies the strict single crossing property (see also Milgrom and Shannon (1994)). Suppose (r_1^*, s_1^*) is the optimal contract at μ_1. At a higher $\mu_2 > \mu_1$, the single crossing property implies that

$$\hat{\lambda}_2 > \lambda_1^*$$

where λ_1^* denotes the equilibrium regulation at μ_1 and $\Pi(r_1^*, s_1^*; \mu_2) = \Pi(1, \frac{\hat{\lambda}_2}{\hat{\lambda}_2 - \sigma}; \mu_2)$. By Theorem 2 in Edlin and Shannon (1998b), r^* is increasing in μ, where the investors' IR

condition is the function G in the theorem. Hence, the new optimum (r_2^*, s_2^*) on the same IR satisfies

$$r_2^* > r_1^* \qquad \text{and} \qquad s_2^* < s_1^*$$

Applying the single crossing property again, we obtain

$$\lambda_2^* > \lambda_1^*,$$

where $\Pi(r_2^*, s_2^*; \mu_2) = \Pi(1, \frac{\lambda_2^*}{\lambda_2^* - \sigma}; \mu_2)$.

The next result we want to establish is that s^u is shifting downward for a higher μ. To do so, for a given r, we determine the sign of $\frac{\partial s}{\partial \mu}$. Suppose the IR is a function of s and μ and compute the total derivative of s with respect to μ.

$$\begin{aligned}
0 &= -u(r)f(\hat{R} - \mu)\frac{r}{s^2}\frac{\partial s}{\partial \mu} + \frac{\partial s}{\partial \mu}\int_0^{\hat{R}} u'(sR)Rf(R-\mu)dR \\
&- \int_0^{\hat{R}} u(sR)f'(R-\mu)dR + u(r)f(\hat{R}-\mu)\frac{r}{s^2}\frac{\partial s}{\partial \mu} + u(r)f(\hat{R}-\mu) \\
&= \frac{\partial s}{\partial \mu}\int_0^{\hat{R}} u'(sR)Rf^{\mu}(R)dR + \int_0^{\hat{R}} u'(sR)sf^{\mu}(R)dR
\end{aligned}$$

Therefore, the partial derivative is

$$\frac{\partial s}{\partial \mu} = -\frac{\int_0^{\hat{R}} u'(sR)sf^{\mu}(R)dR}{\int_0^{\hat{R}} u'(sR)Rf^{\mu}(R)dR}, \tag{30}$$

which is negative. This result is intuitive. If the return prospect is improved unambiguously, an investor is willing to lend more to draw the same utility.

Therefore, the constraint on the profit maximization problem becomes looser and the banks can raise their expected profit under the new IR condition. That is, at the unique contract $(\check{r}_2, \check{s}_2)$ and μ_2 considering both effects, we have

$$\Pi(\check{r}_2, \check{s}_2; \mu_2) > \Pi(r_2^*, s_2^*; \mu_2)$$

Since $\Pi(\check{r}_2, \check{s}_2; \mu_2) = \Pi(1, \frac{\check{\lambda}_2}{\check{\lambda}_2 - \sigma}; \mu_2)$ and $\Pi(r_2^*, s_2^*; \mu_2) = \Pi(1, \frac{\lambda_2^*}{\lambda_2^* - \sigma}; \mu_2)$, it follows that

$$\check{\lambda}_2 > \lambda_2^*,$$

the result we seek.

Proof of Proposition 3 **Claim 1.** From the proof of Proposition 2, we see $r'(\mu) > 0$ and $s'(\mu) < 0$, delivering the first claim.

Claims 2 and 3. The second and third claims are proven in the following three steps.

Step 1: Change in Profit

I drop the superscript μ from the distribution F because it causes no confusion. As

before, SB's profit is given by

$$\Pi(r,s;\mu) = \frac{s}{s-1}\int_{r/s}^{\infty} RdF - \frac{r}{s-1}\int_{r/s}^{\infty} dF$$

and its solvency probability

$$P(r,s;\mu) = \overline{F}\left(\frac{r}{s}\right) = \int_{r/s}^{\infty} f^{\mu}(R)dR$$

Note

$$\begin{aligned}
\frac{\partial P}{\partial \mu} &= -\int_{r/s}^{\infty} f^{\mu\prime}dR = f^{\mu}\left(\frac{r}{s}\right)\\
\frac{\partial P}{\partial r} &= -\frac{1}{s}f^{\mu}\left(\frac{r}{s}\right)\\
\frac{\partial P}{\partial s} &= \frac{r}{s^2}f^{\mu}\left(\frac{r}{s}\right)
\end{aligned}$$

Next, the partial derivatives of $\Pi(r,s;\mu)$ are given by

$$\begin{aligned}
\frac{\partial \Pi}{\partial \mu} &= -\frac{s}{s-1}\int_{r/s}^{\infty} Rf^{\mu\prime}dR + \frac{r}{s-1}\int_{r/s}^{\infty} f^{\mu\prime}dR\\
&= \frac{s}{s-1}\frac{r}{s}f^{\mu}\left(\frac{r}{s}\right) + \frac{s}{s-1}\int_{r/s}^{\infty} f^{\mu}dR - \frac{r}{s-1}f^{\mu}\left(\frac{r}{s}\right)\\
&= \frac{s}{s-1}\overline{F}\left(\frac{r}{s}\right)\\
\frac{\partial \Pi}{\partial r} &= -\frac{1}{s-1}\overline{F}\left(\frac{r}{s}\right)\\
\frac{\partial \Pi}{\partial s} &= -\frac{1}{(s-1)^2}\int_{r/s}^{\infty} RdF + \frac{r}{(s-1)^2}\overline{F}\left(\frac{r}{s}\right)
\end{aligned}$$

Suppose μ changes from μ_0 to μ_1, $\Delta\mu \equiv \mu_1 - \mu_0$. Without loss of generality, we focus on the case that $\Delta\mu > 0$. The corresponding contracts are $(r_0, s_0) \equiv (r(\mu_0), s(\mu_0))$ and $(r_1, s_1) \equiv (r(\mu_1), s(\mu_1))$, respectively. First, think of how much the IR curve moves when μ changes for a given fixed r. Using (30), we derive

$$\begin{aligned}
\frac{\Delta s|_{r=r_0}}{\Delta\mu} &= -\frac{\int_0^{r_0/s_0} u'(s_0R)s_0f^{\mu_0}(R)dR}{\int_0^{r_0/s_0} u'(s_0R)Rf^{\mu_0}(R)dR}\\
&< -\frac{s_0\int_0^{r_0/s_0} u'(s_0R)f^{\mu_0}(R)dR}{(r_0/s_0)\int_0^{r_0/s_0} u'(s_0R)f^{\mu_0}(R)dR} = -\frac{s_0^2}{r_0}.
\end{aligned}$$

Rearranging the inequality, we have

$$\frac{r_0}{s_0^2}\Delta s|_{r=r_0} + \Delta\mu < 0 \tag{31}$$

That is, if we look at the values of s on the two IRs for a given r, the difference of s ($< |s_1 - s_0|$) should satisfies the inequality above for the change in μ.

Now we are ready to prove the third claim for shadow banks ($P^{SB} = \overline{F}\left(\frac{r_0}{s_0}\right)$). Let

us compare the solvency probabilities on the two points considered above. The two points share the same interest rate r_0, but μ and s are different. Thus, from the partial derivatives, the change in SB's solvency probability between the two points is given by,

$$\begin{aligned}\Delta P^{SB}|_{r=r_0} &= \frac{\partial P^{SB}}{\partial s}\Delta s|_{r=r_0} + \frac{\partial P^{SB}}{\partial \mu}\Delta\mu \\ &= \frac{r_0}{s_0^2} f^{\mu}\left(\frac{r_0}{s_0}\right)\Delta s|_{r=r_0} + f^{\mu}\left(\frac{r_0}{s_0}\right)\Delta\mu \\ &= f^{\mu}\left(\frac{r_0}{s_0}\right)\left(\frac{r_0}{s_0^2}\Delta s|_{r=r_0} + \Delta\mu\right) < 0,\end{aligned}$$

by (31). So the default probability goes up when r is fixed at $r = r_0$. If we remove this restriction that $r = r_0$, the default probability increases even more because r becomes bigger and s becomes smaller along the IR curve at $\mu = \mu_1$. We conclude that the default probability of the shadow banks increases when μ increases.

To prove the remaining parts, we turn our attention to the profit function. What is the change in profit induced by the change in μ, denoted by $\Delta\Pi$? First, note that, with $\Delta r = 0$,

$$\Delta\Pi^{res} \equiv \left[-\frac{1}{(s_0-1)^2}\int_{r_0/s_0}^{\infty} RdF + \frac{r_0}{(s_0-1)^2}\overline{F}\left(\frac{r_0}{s_0}\right)\right]\Delta s + \frac{s_0}{s_0-1}\overline{F}\left(\frac{r_0}{s_0}\right)\Delta\mu, \quad (32)$$

from the partial derivatives of $\Pi(r, s)$. The superscript res signifies 'restricted' to $\Delta r = 0$. Since a SB maximizes its profit on the IR curve, the change in Π when r and s freely move on the new IR should be bigger than that when r is fixed. Thus,

$$\begin{aligned}\Delta\Pi &\geq \Delta\Pi^{res} \\ &= \left[-\frac{1}{(s_0-1)^2}\int_{r_0/s_0}^{\infty} RdF + \frac{r_0}{(s_0-1)^2}\overline{F}\left(\frac{r_0}{s_0}\right)\right]\Delta s + \frac{s_0}{s_0-1}\overline{F}\left(\frac{r_0}{s_0}\right)\Delta\mu \\ &> \left[-\frac{1}{(s_0-1)^2}\int_{r_0/s_0}^{\infty} RdF + \frac{r_0}{(s_0-1)^2}\overline{F}\left(\frac{r_0}{s_0}\right)\right]\left[-\frac{s_0^2}{r_0}\Delta\mu\right] + \frac{s_0}{s_0-1}\overline{F}\left(\frac{r_0}{s_0}\right)\Delta\mu \\ &= \frac{s_0}{s_0-1}\frac{1}{r_0}\left[\frac{s_0}{s_0-1}\int_{r_0/s_0}^{\infty} RdF - \frac{r_0}{s_0-1}\overline{F}\left(\frac{r_0}{s_0}\right)\right]\Delta\mu \\ &= \frac{s_0}{s_0-1}\frac{1}{r_0}\Pi\Delta\mu.\end{aligned}$$

This inequality gives a lower bound for the change in profit generated by the increase in μ, $\Delta\mu$. Since this lower bound is positive, we just proved the second claim in the proposition – remind that the SB's profit is the same as that of CB. Note that we use the fact that

$$\frac{1}{(s-1)^2}\int_{r/s}^{\infty} RdF - \frac{r}{(s-1)^2}\overline{F}\left(\frac{r}{s}\right) = \frac{1}{(s-1)^2}\int_{r/s}^{\infty}(R-r)dF$$

is positive, as shown in the proof of Lemma 4.

Step 2: Leverage Comparison

The next step is to show that, under some condition,

$$\frac{s}{s-1}\frac{1}{r} \geq \lambda.$$

To save notations, define a function

$$\varphi(x) \equiv \mathbb{E}[\tilde{R}|\tilde{R} \geq x] = \frac{\int_x^\infty Rf^\mu(R)dR}{\int_x^\infty f^\mu(R)dR}.$$

It is useful to study some properties of $\varphi(x)$. First,

$$\varphi(0) = \mathbb{E}[\tilde{R}] = \mu > 1$$

Second, if there exists an upper bound of the support of R, denoted by R_{sup},[28]

$$\lim_{x \to R_{sup}} \varphi(x) = R_{sup} < 1 + R_{sup}.$$

Hence, a function $\varphi(x) - x - 1$ is positive at $x = 0$ and negative as approaching to R_{sup}. By continuity of $\varphi(\cdot)$, we conclude

Lemma 12. *There exists $\hat{x} \in \mathfrak{R}_+$ such that*

$$\varphi(\hat{x}) = 1 + \hat{x}.$$

Third, the first derivative of φ with respect to x is

$$\begin{aligned}
\varphi'(x) &= \frac{-xf^\mu(x)\int_x^\infty f^\mu(R)dR + f^\mu(x)\int_x^\infty Rf^\mu(R)dR}{[\int_x^\infty f^\mu(R)dR]^2} \\
&= \frac{-xf^\mu(x)\overline{F}(x) + f^\mu(x)R(F(R)-1)|_x^\infty - f^\mu(x)\int_x^\infty (F(R)-1)dR}{\overline{F}(x)^2} \\
&= \frac{-xf^\mu(x)\overline{F}(x) + xf\mu(x)\overline{F}(x) + f^\mu(x)\int_x^\infty \overline{F}(R)dR}{\overline{F}(x)^2} \\
&= \frac{f^\mu(x)\int_x^\infty \overline{F}(R)dR}{\overline{F}(x)^2} < 1,
\end{aligned}$$

by Assumption 2. Combining Lemma 12 with $\varphi'(x) < 1$, we have

$$\varphi(x) < 1 + x$$

for all $x > \hat{x}$. Empirically $\hat{x}$ is well below 1. For example, $\varphi(1)$ is the return conditional on making no loss, which is expected to be significantly smaller than 2.

[28]The only use of this assumption is to make easy the construction of Lemma 12. We can build the lemma more generally without this realistic assumption. However, it is hard to imagine that any asset in the real world has an infinite support of its return distribution over finite horizon.

To proceed, we make a mild assumption:

Assumption 3. *For all values of μ we consider,*

$$\frac{r(\mu)}{s(\mu)} > \hat{x}.$$

It implies that the default return of shadow banks satisfies the inequality above. Since the default return of shadow banks is bigger than that of commercial banks ($\frac{r}{s} > \frac{\lambda-1}{\lambda}$), this assumption is satisfied if $\hat{x}$ is sufficiently small and/or the equilibrium λ is sufficiently bigger than 1.
Assumption 3 implies that

$$0 < \mathbb{E}\left[\tilde{R}\middle|\tilde{R} \geq \frac{r_0}{s_0}\right] - r_0 = \varphi\left(\frac{r_0}{s_0}\right) - r_0 < 1 + \frac{r_0}{s_0} - r_0.$$

This inequality has a direct implication on MRS^{Π} at (r_0, s_0):

$$\begin{aligned} MRS^{\Pi} &= \frac{s_0 - 1}{\mathbb{E}[\tilde{R}|\tilde{R} \geq r_0/s_0] - r_0} \\ &= \frac{s_0 - 1}{\varphi(r_0/s_0) - r_0} \\ &> \frac{s_0 - 1}{r_0/s_0 + 1 - r_0} > 0, \end{aligned}$$

On the (r, s)-plane, the line tangent on the iso-profit curve at (r_0, s_0) is given by

$$s = -MRS^{\Pi}(r - r_0) + s_0$$

Evaluating the value of s on this straight line at $r = 1 (< r_0)$, we obtain

$$\begin{aligned} s &= -MRS^{\Pi}(1 - r_0) + s_0 \\ &> \frac{s_0(s_0 - 1)(r_0 - 1)}{r_0 - (r_0 - 1)s_0} + s_0 \\ &= \frac{s_0}{r_0 - (r_0 - 1)s_0}. \end{aligned}$$

Since the iso-profit curve is convex on the (r, s)-plane and the value of s on the iso-profit curve passing (r_0, s_0) at $r = 1$ is $\frac{\lambda_0}{\lambda_0 - 1}$, we have

$$\frac{\lambda_0}{\lambda_0 - 1} > \frac{s_0}{r_0 - (r_0 - 1)s_0},$$

implying

$$\lambda_0 < \frac{s_0}{r_0(s_0 - 1)}, \tag{33}$$

the result we want to prove. Note that the subscript 0 has no meaning for this result.

Step 3: Higher Default Probability

In equilibrium, the profits of SB and CB are equal, so the change in profit is also the same for both types of bank. Thus the change in CB's profit resulting from the change in μ also satisfies

$$\begin{aligned}\Delta\Pi &\geq \frac{s_0}{s_0-1}\frac{1}{r_0}\Pi\Delta\mu \\ &> \lambda_0\Pi\Delta\mu,\end{aligned}$$

where the inequality comes from (33).
Next, we derive the implication of this inequality that the CB's profit function satisfies. Since r is fixed at 1 for CB, this part is essentially an application of what we showed in the first step to a special case. In the CB's case, we have $\lambda = \frac{s}{s-1}$ and $r = 1$. So, in order to obtain the expression for $\Delta\Pi$ for CBs, we can use (32) without res:

$$\Delta\Pi = \left[-(\lambda_0-1)^2\int_{\frac{\lambda_0-1}{\lambda_0}}^{\infty} RdF + (\lambda_0-1)^2\overline{F}\left(\frac{\lambda_0-1}{\lambda_0}\right)\right]\Delta s + \lambda_0\overline{F}\left(\frac{\lambda_0-1}{\lambda_0}\right)\Delta\mu$$

Imposing the inequality that the change in CB's profit should satisfy, we have

$$\left[-(\lambda_0-1)^2\int_{\frac{\lambda_0-1}{\lambda_0}}^{\infty} RdF + (\lambda_0-1)^2\overline{F}\left(\frac{\lambda_0-1}{\lambda_0}\right)\right]\Delta s + \lambda_0\overline{F}\left(\frac{\lambda_0-1}{\lambda_0}\right)\Delta\mu > \lambda_0\Pi\Delta\mu$$

Substituting the expression of CB's profit with Π, we obtain

$$\begin{aligned}&\left[-(\lambda_0-1)^2\int_{\frac{\lambda_0-1}{\lambda_0}}^{\infty} RdF + (\lambda_0-1)^2\overline{F}\left(\frac{\lambda_0-1}{\lambda_0}\right)\right]\Delta s + \lambda_0\overline{F}\left(\frac{\lambda_0-1}{\lambda_0}\right)\Delta\mu \\ &\qquad > \lambda_0\left[\lambda_0\int_{\frac{\lambda_0-1}{\lambda_0}}^{\infty} RdF - (\lambda_0-1)\overline{F}\left(\frac{\lambda_0-1}{\lambda_0}\right)\right]\Delta\mu\end{aligned}$$

As noted above,

$$\int_{\frac{\lambda_0-1}{\lambda_0}}^{\infty} RdF - \overline{F}\left(\frac{\lambda_0-1}{\lambda_0}\right)$$

is positive, so the inequality above is simplified to

$$\left(\frac{\lambda_0-1}{\lambda_0}\right)^2\Delta s + \Delta\mu < 0.$$

This inequality is a special case of (31) for CB. To interpret this relation, let's compute the change in CB's solvency probability ($P^{CB} = \overline{F}\left(\frac{\lambda-1}{\lambda}\right)$) by $\Delta\mu$. We already gave the partial derivatives of the solvency probability in the first step. Using the results, we

have

$$\begin{aligned}
\Delta P^{CB} &= \frac{\partial P^{CB}}{\partial s}\Delta s + \frac{\partial P^{CB}}{\partial \mu}\Delta \mu \\
&= \left(\frac{\lambda_0 - 1}{\lambda_0}\right)^2 f^{\mu_0}\left(\frac{\lambda_0 - 1}{\lambda_0}\right)\Delta s + f^{\mu_0}\left(\frac{\lambda_0 - 1}{\lambda_0}\right)\Delta \mu \\
&= f^{\mu_0}\left(\frac{\lambda_0 - 1}{\lambda_0}\right)\left[\left(\frac{\lambda_0 - 1}{\lambda_0}\right)^2 \Delta s + \Delta \mu\right] < 0,
\end{aligned}$$

by the inequality above. We conclude that an increase in μ lowers the CB's solvency probability and, in turn, raises the CB's default probability.

Claim 4. Differentiating the solvency probability of CB with respect to μ, we obtain

$$\begin{aligned}
\frac{dP^{CB}}{d\mu} &= \frac{d}{d\mu}\int_{\frac{\lambda-1}{\lambda}}^{\infty} f(R-\mu)dR \\
&= -\frac{\lambda'(\mu)}{\lambda(\mu)^2} f^{\mu}\left(\frac{\lambda - 1}{\lambda}\right) - \int_{\frac{\lambda-1}{\lambda}}^{\infty} f'(R-\mu)dR \\
&= -\left(\frac{\lambda'(\mu)}{\lambda(\mu)^2} - 1\right) f^{\mu}\left(\frac{\lambda - 1}{\lambda}\right).
\end{aligned}$$

From the last claim, we know that this total derivative is negative, implying $\lambda'(\mu) > \lambda(\mu)^2$.
The per-capital expected cost of deposit insurance, denoted by $T(\mu)$, is given by

$$\begin{aligned}
T(\mu) &= \mathbb{E}^{\mu}\left[\max\{0, \lambda(\mu) - 1 - \lambda(\mu)\tilde{R}\}\right] \\
&= \lambda(\mu)\int_0^{\frac{\lambda-1}{\lambda}} \left(1 - \frac{1}{\lambda(\mu)} - R\right) f^{\mu}(R)dR
\end{aligned}$$

The first derivative is

$$\begin{aligned}
T'(\mu) &= \lambda'(\mu)\int_0^{\frac{\lambda-1}{\lambda}} \left(1 - \frac{1}{\lambda(\mu)} - R\right) f(R-\mu)dR \\
&+ \lambda(\mu)\int_0^{\frac{\lambda-1}{\lambda}} \frac{\lambda'(\mu)}{\lambda(\mu)^2} f(R-\mu)dR \\
&- \lambda(\mu)\int_0^{\frac{\lambda-1}{\lambda}} \left(1 - \frac{1}{\lambda(\mu)} - R\right) f'(R-\mu)dR \\
&= \lambda'(\mu)\int_0^{\frac{\lambda-1}{\lambda}} \left(1 - \frac{\lambda(\mu)}{\lambda'(\mu)} - R\right) dF
\end{aligned}$$

The smallest value of the integrand is positive if and only if $\lambda' > \lambda^2$, a sufficient condition for $T'(\mu) > 0$. Thus, the last claim is sufficient for this claim.

References

[1] Acharya, Viral, and Matthew Richardson, 2009, Causes of the Financial Crisis, *Critical Review* 21(2-3), 195-210.

[2] Acharya, Viral, Philipp Schnabl, and Gustavo Suarez, 2011, Securitization Without Risk Transfer, *Journal of Financial Economics*, forthcoming.

[3] Adrian, Tobias, and Hyun Song Shin, 2010, Liquidity and Leverage, *Journal of Financial Intermediation* 19, 418-437.

[4] Aiyar, Shekhar, Charles W. Calomiris, and Tomasz Wieladek, 2012, Does Macro-Pru Leak? Evidence from a UK Policy Experiment, *NBER Working Paper Series, Working Paper 17822.*

[5] Coval, Joshua, Jakub Jurek, and Erik Stafford, 2009, The Economics of Structured Finance, *Journal of Economic Perspectives* 23, 3-25.

[6] Edlin, Aaron, and Chris Shannon, 1998, Strict Single Crossing and the Spence-Mirrlees Condition: A Comment on Monotone Comparative Statics, *Econometrica* 66, 1417-1425.

[7] Edlin, Aaron, and Chris Shannon, 1998, Strict Monotonicity in Comparative Statics, *Journal of Economics Theory* 81, 201-219.

[8] Goodhart, Charles, Anil Kashyap, Dimitrios Tsomocos, and Alexandros Vardoulakis, 2011, Financial Regulation in General Equilibrium, *Working Paper.*

[9] Gorton, Gary, and Andrew Metrick, 2010, Regulating the Shadow Banking System, *Brookings Papers on Economic Activity.*

[10] Hanson, Samuel, Anil Kashyap, and Jeremy Stein, 2011, A Macroprudential Approach to Financial Regulation, *Journal of Economic Perspectives* 25, 3-28.

[11] Kashyap, Anil, and Jeremy Stein, 2004, Cyclical implications of the Basel-II capital standards, Federal Reserve Bank of Chicago, *Economic Perspectives* 28, 1831.

[12] Milgrom, Paul, and Chris Shannon, 1994, Monotone Comparative Statics, *Econometrica* 64, 157-181.

[13] Pozsar, Zoltan, Tobias Adrian, Adam B. Ashcraft, and Haley Boeskey, 2010, Shadow Banking, *Federal Reserve Bank of New York Staff Report 458.*

[14] Shift for Goldman and Morgan Marks the End of an Era, *NY Times, Sep 2008*, http://www.nytimes.com/2008/09/22/business/22bank.html.

[15] Taylor, John B., 1993, Discretion Versus Policy Rules in Practice, *Carnegie-Rochester Conference Series on Public Policy* 39, 195-214.