# Optimal Capital Regulation with Two Banking Sectors

Taejin Kim and Vishal Mangla *

November 17, 2011

**Abstract**

We present the case for pro-cyclical capital regulation policy as opposed to the generally accepted counter-cyclical policy in the current literature. Our argument is based on the fact that banks move around their capital in and out of the regulation umbrella depending on 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 counter-cyclical policy) forces the banks to move to the shadow banking sector where they take too much risk. Therefore, the policy aimed at controlling the systemic risk during booms should incentivize the banks to be regulated by relaxing the regulation. These forces are reversed during busts. In our model, the optimal capital regulation policy is specified as a schedule of a macro variable: relative banking sector size. Under this specification, a feedback loop exists between the policy instrument and the macro variable. This has the effect that the economy transitions smoothly from one equilibrium to another as the underlying fundamentals change. Our analysis is built around characterizing the solution of a contracting problem between a shadow (unregulated) bank and its investors.

*Northwestern University. We are deeply indebted to Arvind Krishnamurthy for his guidance. We also thank seminar participants at Northwestern University for helpful comments. All errors are our own. E-mails: {taejin-kim, v-mangla}@northwestern.edu.

# 1 Introduction

In this paper, we look for the capital regulation policy that achieves the optimal systemic risk exposure when banks can choose to be unregulated. Our work is motivated by the existence of regulatory arbitrage in the financial intermediation sector before and during the crisis of 2007-2009. Before the crisis, we saw a burgeoning of the shadow banking sector (Coval, Jurek and Stafford (2009), Gorton and Metrick (2010)). Figure 1 (reproduced from Pozsar et al. (2010)) illustrates that the sector size doubled from \$10 trillion in the year 2000 to \$20 trillion at its peak in March 2008. An important component of this is the widespread growth of SIVs & SPVs by means of which the commercial banks take their investments off balance sheet outside the regulation umbrella (Acharya, Schnabl, and Suarez (2010), Figures 5 & 7 in Gorton and Metrick (2010)). We argue that this excessive flow of investment outside the regulated banking sector was the result of 'too tight' regulation. During the crisis, however, we witnessed a few investment banks acquiring the status of bank holding company (*NY Times*, Sep 2009). More generally, a number of hitherto unregulated sections of the intermediary sector accepted the government support. This flow of investment under the regulation umbrella was the result of 'too loose' regulation during the crisis.

To model this strategic behavior of financial institutions, we consider a setting where banks can choose their type as one of regulated (commercial) or unregulated (shadow) depending on the current status of the (capital requirement) regulation. In making this decision, the banks compare the cost of being regulated – limit on risk exposure – with the benefit – access to cheaper funding. Knowing this behavior of the banks, the government chooses a regulation that achieves the 'right' mix of the commercial and the shadow banking sectors in the economy. The most interesting outcome of our analysis is the pro-cyclical nature of the capital regulation policy – capital requirement is loose during good times and tight during bad times. This is completely opposite to the current stance on the capital requirement regulation, which argues for counter-cyclical regulation policy (example Kashyap and Stein (2004)). We attribute this difference to the existing literature ignoring the ability of banks to move around their capital in and out of the regulation umbrella.

The current literature proposes a higher capital requirement during good times which can act as buffer against losses during bad times. However, this advice ignores the fact that banks that are averse to raising costly capital will simply switch their types from regulated to unregulated (move their investment outside the regulator's purview as they did before the crisis). This will in turn result in a too big shadow banking sector and the correspondent high systemic risk. To limit the risk in good times, we propose that the government actually requires a lower

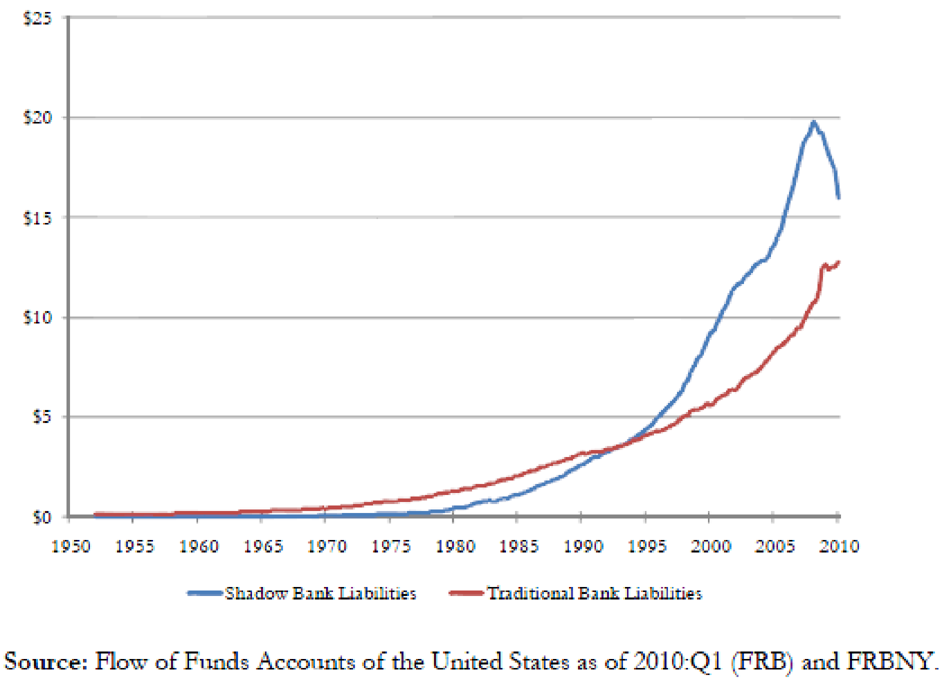


Figure 1: Shadow Bank Liabilities vs. Traditional Bank Liabilities, \$ trillion (Pozsar et al. (2010))

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/fire sell in order to maintain their capital ratio. However, too low capital requirement would mean all banks would like to be regulated since they receive government's funding support at a low cost. So, there will be too many commercial banks and the economy-wide investment profile will 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 systemic risk during bad times[1].

Another important feature of our analysis is that we pursue a regulation policy that is market-based. We will see that under the optimal policy, the capital requirement $\lambda$ varies according to some pre-specified time-invariant rule: $\lambda = \Omega(\xi)$. The schedule $\Omega$ is independent of the economic fundamentals and is increasing in $\xi$ ($\Omega' > 0$). $\xi$ is the relative size of the shadow banking sector with respect to that of

[1]In our baseline model, we ignore the traditional forces that generate counter-cyclical regulation. However, in section 10 we consider an extension of our model where even when we introduce these traditional forces, the pro-cyclical regulation result survives.

the commercial banking sector[2]. We expect that the level of capital requirement $\lambda$ impacts the relative sector size $\xi$. However, under this policy specification, the reverse causation also holds, $\xi$ impacts $\lambda$. Due to this feedback loop, once the policy schedule has been announced ('etched in stone'), the market self-adjusts to attain the optimal risk exposure as the underlying fundamentals change[3]. In this sense, the policy is also robust to the business cycle fluctuations. A feedback loop between the policy instrument and the macro variable being targeted is also a feature of Taylor rule 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. Also, we show that monotonicity of $\Omega$ makes the policy robust to small measurement errors on the part of the government in that small errors lead to only a small deviation in the risk exposure from the optimum. In addition to its own appeal, this robustness feature also delivers the uniqueness of optimal regulation in a simple setup.

The capital requirement regulation considered in this paper takes the following rule: the standard deviation of a bank's asset portfolio return is upper bounded by the scalar $\lambda$ (the policy instrument; same for all banks) times the bank's equity capital. This form is in accordance with Basel II in which the banks need to hold capital commensurate with the risk of their investments. We show that the optimal policy instrument $\lambda$ varies with the underlying economic fundamentals in a pro-cyclical manner. Even though we consider capital requirement as the policy instrument in our analysis, we would arrive at the same two key features – pro-cyclicality and robustness if we use deposit insurance premium that is increasing in the risk of bank's asset portfolio as the policy instrument.

In our economy, there are two assets available for investment – a riskless asset and a risky asset. There are three agents – a government, banks and investors. While the banks are risk-neutral, the investors are risk-averse. The investors are unable to access the asset market themselves and invest in the banks' securities. The banks can access the asset market. If they are unregulated, the banks invest only in the risky asset. The benefit of being regulated is cheaper financing due to deposit insurance (we assume zero fee for deposit insurance). The cost of being regulated is a government imposed limit on the maximum allowable risk of the bank's asset portfolio. 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 asset (since banks are risk neutral) and there is too much

[2]We will see $\xi$ is closely related to the systemic risk exposure. For any given value of fundamentals, there is an optimal/target value of $\xi$ that the government is able to achieve using the policy $\Omega$.

[3]Committing to the schedule $\Omega$ to implicitly obtain the optimal $\xi$ rather than directly managing $\xi$ is analogous to managing prices (and not quantities) in the terminology of Weitzman (1974).

systemic risk[4]. The government's objective is to obtain the optimal systemic risk exposure. The government, while deciding the capital requirement, takes into account the banks' incentive to become regulated. It is this consideration on the government's part that distinguishes our analysis from that of the current literature.

There are three main empirical predictions of our model: (i) the relative size of the shadow banking sector with respect to that of the commercial banking sector is pro-cyclical when the capital requirement is held fixed; (ii) the leverage of the shadow banking sector is pro-cyclical; and (iii) shadow banks offer pro-cyclical interest rate to their investors. The first two predictions find support in data.

It is also important to note that even when the form of regulation considered in this paper is individual bank specific (that is, the regulation specifies the maximum risk that each bank can assume), the regulation actually controls economy-wide/systemic risk. This is because there is only one risky asset in this economy (the market portfolio) and thus all banks' investments are perfectly correlated. This is somewhat similar to the effect achieved by Gennaioli, Shleifer and Vishny (2011) where banks, in order to satisfy the investors high demand for safe debt, pool and tranch respective loan portfolios to diversify away idiosyncratic risk and concentrate systemic risk. Also, in this sense, this paper is related to the literature on macro-prudential regulation. Our focus is closer to aggregate risk-shifting incentive of Acharya (2009) compared to aggregate balance sheet shrinkage considered in Hanson, Kashyap, and Stein (2011).

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 bring the entire shadow banking sector under the regulation umbrella. Irrespective of the outreach of the regulation, there will always exist some banks in the shadows (due to regulatory arbitrage, for example). That being the case, the best regulators can do is to implicitly influence (and not overtake) the shadow banking sector by designing meaningful regulation[5]. 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

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

[5]Interestingly, for some values of the economic fundamentals, we show that one equilibrium in our model is a loose enough regulation that all banks find it profitable to be regulated. An immediate disadvantage of having all the banks regulated is huge deposit insurance cost bore by the government (for simplicity, we ignore this cost in our analysis).

advocated (that is, choose the 'right policy instrument' rather than the 'right level' of a given policy instrument).

Finally, we analyze optimal regulation in a financial crisis. One of the notable features of a crisis is a 'flight to quality'. In our model, this means that the investors put their investment in the commercial banks during a crisis. One simple way to achieve this is to define a crisis as a transitory jump in the investors' risk aversion. The government's objective is to obtain the right systemic risk exposure even when it cannot track the investors' risk aversion during the crisis. This can be achieved if we assume that the duration of the crisis is short so that the fundamentals do not vary much during the crisis. Under that assumption, we derive a policy schedule robust to fluctuation in investors' risk aversion in much the same way as we derive the optimal policy schedule in the main analysis.

# 2 Setup

Time is continuous, $t \in [0, 1]$. The final date is normalized to 1.

## 2.1 Assets

There is a riskless asset and a risky asset (best thought of as the asset with the highest Sharpe ratio; the market portfolio). The riskless asset returns 1 and the risky asset returns $\tilde{R}$ at $t = 1$. The distribution of $\tilde{R}$ at time $t \in [0, 1]$ is $F_t$, with mean $\mu_t$ and standard deviation $\sigma_t$. The density $f_t$ is continuously differentiable as many times as necessary. Note that the asset pays only at the terminal date and the agents' beliefs about it are changing over time. We impose $\mu_t \equiv \mathrm{E}_t[\tilde{R}] > 1$ for all $t$. $F_t$ is completely determined by $\mu_t$ and $\sigma_t$, denoted by $F^{\mu_t,\sigma_t}$. Moreover, all $F^{\mu_t,\sigma_t}$ belong to one location-scale family, like normal distribution. Hence we write $f^{\mu,\sigma} = \frac{1}{\sigma} f(\frac{\tilde{R}-\mu}{\sigma})$ for some reference $f$, for instance, with $\mu = 0$ and $\sigma = 1$. Since the asset pays only at the terminal date and there is no time value of money, all decisions are made in an identical way irrespective of its timing. In other words, if $F_t = F_{t'}$, the economy should look exactly the same for $t$ and $t'$. Despite this static nature, we use a multiple (infinite) time setting because one of our objectives is to come up with a regulation scheme that achieves our goal for any realized path of $\mu_t$ and $\sigma_t$. This setup enables us to avoid bringing in another layer of complexity from dynamic programming while it clearly exhibits how our mechanism delivers the robustness[6].

[6] There is a caveat in this setup: Since the assets pay nothing interim and 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 $t \in [\frac{1}{2}, 1]$. Please keep in mind that we

The assets can be acquired and liquidated at the same price of one at any time. The assets can be thought of investment projects and incur no loss when liquidated.

## 2.2 Banks

Banks of measure 1 are risk-neutral so invest only in the risky asset if there is no 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 so that when they are unregulated, the banks undertake more risks than socially optimal (see regulator's objective below). We assume that the investors are unable to monitor or enforce the portfolio choice of the banks, so it is not contractible. As will be discussed later, the contract is on how the total return will be distributed between the bank and the investor.

Each banker has her own equity capital, $k$ which is distributed according to $G$. So the entire wealth owned by the banks is $\int kdG$. Except for the size of their own capital, the banks are identical. They have the same information and investment opportunity. We assume no heterogeneity in their skills or effort levels.

Therefore, if a bank raises $d$ from investors, the (gross) return of its portfolio realized at $t = 1$ is

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

where $\beta$ is the weight on the risky asset and $0 \leq \beta \leq 1$.

The key feature of this paper is that the banks can choose whether to be regulated. We call the regulated banks *commercial banks, CB* while we use *shadow banks, SB* for the unregulated banks[7]. In our model, SBs set $\beta = 1$ and CB set $\beta$ equal to the maximum value allowed under the regulation (details to follow). Examples of CBs are Bank of America and Citigroup that take deposits from investors. These deposits are insured by the government. SB sector includes investment banks (Lehman, Goldman Sachs (before crisis) etc.), finance companies, asset-backed commercial paper (ABCP) conduits, limited-purpose finance com-

adopt this setting solely for the purpose to show 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 agents participate; that is, they choose optimal actions at any $t$ for the scenario that 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.

[7]Given that a bank can only be of one type in our model, how do we interpret 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 perpective – what matters is the ratio of unregulated to regulated bank capital in the economy.

panies, structured investment vehicles, credit hedge funds, money market mutual funds, securities lenders.

## 2.3 Investors

There is a continuum of investors with total measure $W$. Each investor has one unit of good to invest and all investors have the same risk-averse preference. Her preference is represented by a Bernoulli function $u(\cdot)$ with $u(1) = 0$, $u' > 0$ and $u'' < 0$. Also, we impose the integrability condition: $\int_0^\infty |u(\tilde{R})|dF < \infty$. The investors do not have direct access to the risky asset and so have to use the banks to earn risky cash flow profiles (limited participation). Each investor is assumed to be atomic in the sense that her unit of good is indivisible. She can choose only one bank for her investment. The bank can either be regulated or unregulated (see below). All commercial (regulated) banks offer the same (risk-free) interest rate to their investors and are thus identical from the investors' perspective.

Since an unregulated bank offers a risky cash flow, a portfolio choice problem naturally arises if we allow an investor to invest in both types of banks. The restriction on the access to banks, therefore, captures the idea that an individual investor does not consider overall riskiness of the total investment of the economy. This is a simplifying assumption that lets us avoid the computational complexity caused by the portfolio choice problem of an individual investor but is not essential for our results.

## 2.4 Regulator

The case for regulation in our model is motivated as follows: The investors are the ultimate owners of the banks. But due to lack of expertise, the investors have to provide the control of investment in the hands of bankers. We impose market incompleteness that the banks choice of portfolio is not contractible[8] (but terms of funding are). With this market incompleteness, we depart from the assumptions of the first welfare theorem – the banks impose negative externality on the investors by undertaking investments that are riskier than what investors like. Inability of the investors/owners to dictate what portfolio the banks should choose is the market failure in our model.

### 2.4.1 Objective

The main concern of the government is to keep the aggregate risk at a desirable level depending on the economic condition at the time, with the friction that the

[8]Since we have a dispersed investor base rather one aggregate investor who may be a more powerful negotiator, this assumption is easier to defend.

banks as investment vehicles do not care about the risk of their investment. Hence it is natural that the government maximizes the s' utility over the total return from the investment. This objective implicitly assumes that the total cash flows including the banks' profit and the government's insurance payout will be attributed to the investors eventually. Therefore, while we use the investors' preference in the regulator's problem, we consider not only the cash flows to the investors (from the banks) but also those to other agents. Then, it amounts to maximizing the utility over the total return from the investment of the economy. In our model, this is equivalent to choosing the fraction of aggregate wealth invested into the risky asset: If, on aggregate, $\beta$ fraction of total wealth is invested into the risky asset, the aggregate return is $\tilde{X}(\beta) = 1 + \beta(\tilde{R} - 1)$. Then, the government chooses $\beta$ that maximizes $\mathbb{E}u(\tilde{X}(\beta))$. The solution $\beta^*$ to this problem is some function $\phi(\mu, \sigma)$ (for CRRA($\gamma$) utility function, $\phi(\mu, \sigma) = \frac{\mu - 1}{\gamma\sigma^2}$).

### 2.4.2 Regulation

If the banks choose to be regulated, they are able to receive deposit insurance provided by the government. We assume that the deposit insurance is free[9]. As an exchange, the government imposes regulations on the banks' activities. Here we assume that the government imposes a restriction on the banks' portfolios, which has the specific form:

$$\text{standard deviation of the bank's portfolio return} \leq \lambda k, \tag{2}$$

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. This specification is consistent with 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).

In the current setup, this regulation amounts to a limit on the leverage a commercial bank can take for its investment in the risky asset (see section 4 for details). This makes sense – If a bank invests its deposits in a risk-free asset, the government does not need to regulate bank's leverage because there is no chance of default; the government is only concerned about the fraction of intermediaries' investment in the *risky* asset.

[9]There is little room for the insurance premium to play a role in this paper, due to the following assumptions: i) There is only one risky asset available to all banks, ii) there is no incentive issue among agents, so we do not need to consider the insurance premium as a tool for incentive-based regulation, and iii) we do not explicitly include the expected payout during financial distress in the government's objective function. Note, our results hold for any other government objective that delivers the form of setting the systemic risk exposure equal to a function that is increasing in $\mu$ and decreasing in $\sigma$.

# 3 Contracts

The investors cannot enforce a choice of portfolio by contracts. While the portfolio choice of banks is not contractible, we assume that the total return from bank's investments is verifiable and a contract on return distribution is enforceable. Hence the contract, as common in the literature, specifies each party's payoff for every possible state of the world. Only uncertainty is the return of the risky asset so the contract terms are functions of $\tilde{R}$.

The optimal contract between the investors and the banks is a debt contract. Intuitively, risk aversion of the investors implies that the cheapest way for the banks to compensate the investors is to make risk-less payments to the extent possible. Being risk-neutral themselves, the banks do not care about the riskiness of their own compensation. A debt contract transfers welfare from the banks to the investors most efficiently[10].

We now investigate how the terms of debt contract are determined. We will assume that the banks possess the whole bargaining power and extract the total rent from the investors.

## 3.1 Shadow Banks and Investors

At time $t$, a shadow bank with equity capital $k$ and total asset to liability ratio $s_t(k)$ offer a take-it-or-leave-it interest rate $r_t(k)$ to a potential investor based on the belief $(\mu_t,\ \sigma_t)$ (henceforth, we suppress the time subscript for brevity). An atomic investor accepts the offer and lend her unit of investment to the bank if it is individually rational for her to do so (this depends on $s(k)$; see below). By an argument of rational expectation, in equilibrium, the number of investors that lend this bank is such that the bank's total asset to liability ratio is actually $s(k)$. That is, $\frac{k}{s(k)-1}$ measure of investors write identical contracts with the bank that has capital $k$ - the only term of contract is $r(k)$. Then, equilibrium $r(\cdot)$ and $s(\cdot)$ are solution(s) to the optimization problem:

$$\max_{r(\cdot)>1,s(\cdot)>1} \mathbb{E}\max\left\{\frac{ks(k)}{s(k)-1}\tilde{R}-\frac{k}{s(k)-1}r(k),0\right\} \tag{3}$$

$$\text{s.t. } \mathbb{E}u\left(\min\left\{r(k),s(k)\tilde{R}\right\}\right)=u(1) \tag{4}$$

To see this, notice that the bank's total asset size is $\frac{ks(k)}{s(k)-1}$ and the total return at $t=1$ is $\frac{ks(k)}{s(k)-1}\tilde{R}$. At $t=1$, the bank owes $r(k)$ to each of its $\frac{k}{s(k)-1}$ investors. If the bank is unable to service its debt, it defaults, earns zero itself and distribute

[10] Deposits and asset-backed commercial paper (ABCP) respectively are examples of debt security offered by commercial and shadow banks.

the whole return equally among its investors in which case each investor receives $s(k)\tilde{R}$.

The particular structure of the problem greatly simplifies the analysis. Notice that once we factor out $k$ from the objective function (3) (to obtain expected profit per unit capital), $k$ enters the optimization problem only through $r(k)$ and $s(k)$. This means the problem is identical for all values of $k$. This observation delivers us the simplification:

**Lemma 1.** *Functions $r(\cdot)$ and $s(\cdot)$ are constant functions.*

This result ensures that the shadow banks' expected profit is linear in its equity capital $k$.

In the following, we will show that the solution $(r, s)$ to the contracting problem stated above is unique. But before we are able to do so, we have to characterize the investors' individual rationality (IR) constraint (4) and the banks' profit function (3).

### 3.1.1 Investors' IR Constraint

Given $r$ and $s$ are independent of $k$, condition (4) delivers a relation between $s$ and $r$, denoted by $s^u(r)$. $1/(s^u(r) - 1)$ is the investors' supply curve – amount of funds the investors are willing to supply a bank with unit capital when it offers interest rate $r$.

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

In addition to the lemma above, we can show that either $s$ is bounded away from 1 or $r$ is upper bounded.

**Proposition 1.** *If $\mathbb{E}u(\tilde{R}) > u(1)$, $r$ is upper bounded. If $\mathbb{E}u(\tilde{R}) < u(1)$, $s^u$ is bounded away from 1.*

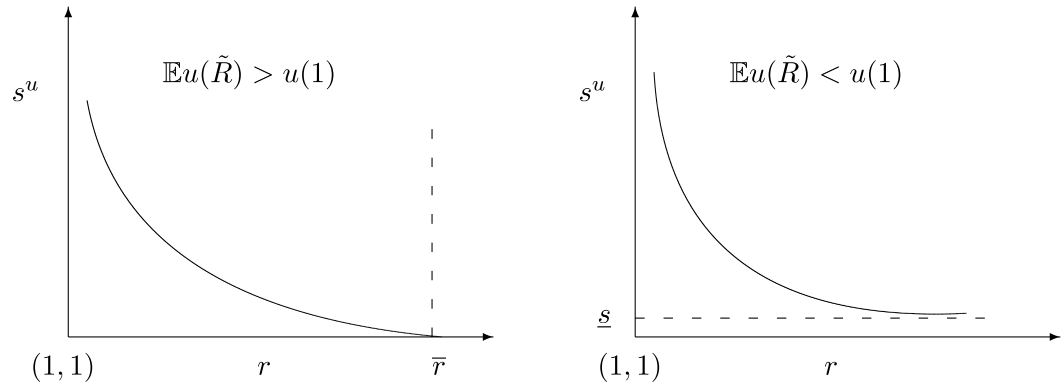


Figure 2: Finite Supply of Debt

This proposition delivers the condition that the investors do not supply infinite amount of money to the shadow banks. If $\mathbb{E}u(\tilde{R}) > u(1)$, the investors are willing to lend infinite amount for a sufficiently large but finite return ($r \geq \overline{r}$). If $\mathbb{E}u(\tilde{R}) < u(1)$, the supply has an upper bound for all $r$. Figure 2 depicts these two cases. From here on, we assume the condition that ensures finite supply of debt[11].

**Assumption 1.** *Finite Supply of Debt*

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

If $u(\cdot)$ represents CRRA($\gamma$) preference, this assumption puts a lower bound on the relative risk aversion coefficient $\gamma$.

### 3.1.2 Shadow Banks' Profit Function

It is convenient to define a new function related to the expected profit of banks. Define a function

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

As we will see shortly, this function computes the per-unit-equity profit of SB when the asset-liability ratio is $y$ and the contract is $x$.

Given $r$ and $s$ are independent of $k$, the profit (3) of the shadow bank is

[11] This assumption, along with the maintained assumption of indivisibility of each investor's good, means that even if the investors had direct access to the risky asset, they would prefer to not invest in it.

rewritten as

$$\Pi^{SB}(\tilde{R};k) = \frac{k}{s-1}\max\{s\tilde{R}-r,0\}$$

The expected profit is

$$\begin{aligned}
\overline{\Pi}^{SB}(k) \equiv \mathbb{E}(\Pi^{SB}(\tilde{R};k)) &= \mathbb{E}\left(\frac{k}{s-1}\max(s\tilde{R}-r,0)\right) \\
&= \frac{k}{s-1}\left(s\cdot\mathbb{E}\left[\tilde{R}\middle|\tilde{R}>\frac{r}{s}\right]-r\right)\left(1-F\left(\frac{r}{s}\right)\right) \\
&= k\Pi(r,s).
\end{aligned}$$

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

**Lemma 3.** *The marginal rate of substitution of the iso-profit function is positive and decreasing in $r$.*

### 3.1.3 Solution to the Contracting Problem

With the characterizations of the two functions, 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

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

**Lemma 4.** *For any Bernoulli function $u$, following programs have the same solutions for $r$.*

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

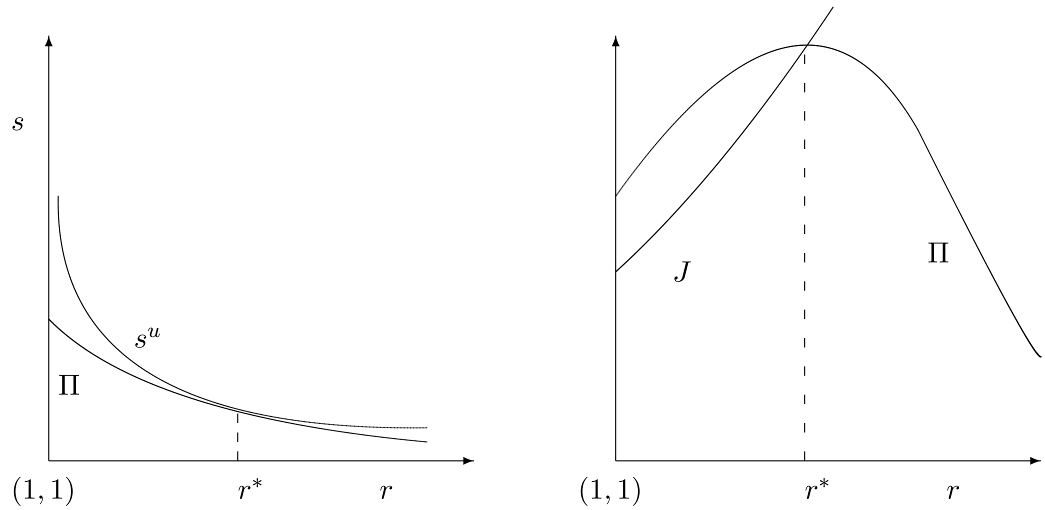


Figure 3: Solution to the Contracting Problem

*Proof.* 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 $MRS^{\Pi}$ 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

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

□

**Corollary 1.** *In the CRRA($\gamma$) case, $J(r) = r^\gamma$.*

Figure 3 displays how the contract term, $r^*$, is determined. The left figure shows that the two marginal rates of substibution coincide, the second statement in Lemma 4. The right figure illustrates the third statement in Lemma 4 that the function $J(r)$ cuts $\Pi(r)$ at the solution. The next lemma proves that the function $J(r)$ is an increasing function in $r$, as in Figure 2. This property will play a crucial role in proving the following proposition.

**Lemma 5.** $J' > 0$, $\lim_{r\to 1} J(r) = 1$, *and* $\lim_{r\to\infty} J(r) = \infty$

**Proposition 2.** *The solution $(r^*, s^*)$ to the contracting problem exists and is unique.*

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

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

We use Proposition 1 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. □

### 3.2 Commercial Banks and Investors

This contract is simple. Due to the deposit insurance, 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 investors 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 commercial banking sector is pinned down by the market clearing condition in the investors' wealth – the commercial banks raise the investors' wealth that is not invested in the shadow banking sector debt.

## 4 Banks' Choice: Commercial vs. Shadow

Given the contracts of the previous section, we now consider the decision of a bank to become regulated (commercial). If a commercial bank holds the portfolio (1), the regulation (2) implies that the bank will invest in the risky asset up to the limit

$$\beta = \frac{\lambda k}{\sigma(k+d)},$$

where $d$ is the amount of debt a CB can raise and is some function $d(k;G)$ of its own capital $k$ and the distribution $G$. The CB's return, $\tilde{R}$ being realized, is

$$\begin{aligned}\Pi^{CB}(\tilde{R};k) &= \max\left\{\left(\frac{\lambda}{\sigma}k\tilde{R} + k + d(k;G) - \frac{\lambda}{\sigma}k\right) - d(k;G), 0\right\} \\ &= k\max\left\{\tilde{R}\frac{\lambda}{\sigma} + 1 - \frac{\lambda}{\sigma}, 0\right\}\end{aligned}$$

Note that the profit of commercial bank does not involve the amount of borrowing from the outside. The regulation in this setting amounts to a restriction on how much risky asset the bank can buy in terms of good (it is $\beta(k + d(k;G)) = \lambda k/\sigma$). Hence, given the regulation, the (risk) leverage ($\lambda/\sigma$) of a commercial bank carries no additional information of the riskiness of the bank, for the bank invests all the good in excess of the limit imposed by the regulation ($k + d(k;G) - \lambda k/\sigma$) into the riskless asset.

A bank decides whether to be SB or CB by comparing the expected values of the payoffs in both scenarios. The expected profit of SB is

$$\overline{\Pi}^{SB}(k) = k\Pi(r,s)$$

whereas

$$\begin{aligned}
\overline{\Pi}^{CB}(k) = \mathrm{E}(\Pi^{CB}(R;k)) &= k\left(\mathrm{E}\left[\tilde{R}\middle|\tilde{R} > \frac{\lambda-\sigma}{\lambda}\right]\frac{\lambda}{\sigma} + 1 - \frac{\lambda}{\sigma}\right)\left(1 - F\left(\frac{\lambda-\sigma}{\lambda}\right)\right) \\
&= k\frac{\lambda-\sigma}{\sigma}\left(\mathrm{E}\left[\tilde{R}\middle|\tilde{R} > \frac{\lambda-\sigma}{\lambda}\right]\frac{\lambda}{\lambda-\sigma} - 1\right)\left(1 - F\left(\frac{\lambda-\sigma}{\lambda}\right)\right) \\
&= k\Pi\left(1, \frac{\lambda}{\lambda-\sigma}\right).
\end{aligned}$$

Both expected profits are linear functions of $k$ with a positive slope and the slope is independent of $k$. Therefore, the bank's decision making is very simple:

$$\begin{array}{ll}
\text{If } \Pi(1, \frac{\lambda}{\lambda-\sigma}; \mu, \sigma) > \Pi(r, s; \mu, \sigma), & \text{CB} \\
\text{If } \Pi(1, \frac{\lambda}{\lambda-\sigma}; \mu, \sigma) < \Pi(r, s; \mu, \sigma), & \text{SB} \\
\text{If } \Pi(1, \frac{\lambda}{\lambda-\sigma}; \mu, \sigma) = \Pi(r, s; \mu, \sigma), & \text{indifferent}
\end{array}$$

If these profit functions were not linear in bank's capital, 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). But this means equilibrium level of risk is indeterminate. 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. The relative sector size is used to estimate the systemic risk exposure (see section 6).

# 5 Timeline

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

1. The government announces the regulation schedule $\lambda(\cdot)$.
2. The agents (investors and banks) observe the fundamental of the economy ($\mu$ and $\sigma$).
3. An equilibrium is reached where no agent has an incentive to deviate: Given the contracts written among agents, i) the shadow banks maximize their profit given the investors' participation constraint and leverage, ii) agents rationally expect a bank's leverage when they write a contract, iii) each bank does no better by changing its type, and iv) the commercial banks abide by

the regulation.

At any interim $t \in (0, 1)$, if $\mu$ or $\sigma$ changes, they repeat 2 and 3 of $t = 0$ procedure. At $t = 1$, all uncertainties are resolved and returns are realized.

# 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 – pro-cyclicality. Second, we relate the relative shadow banking sector size (the macro variable observed by the government) with the aggregate risk exposure. The optimal shadow banking sector size is also found to be pro-cyclical. Third, we tie conditions obtained in the previous two steps to obtain a policy schedule specified only in terms of the variable observed by the government. In particular, the regulation is not be specified in terms of economic fundamentals $(\mu, \sigma)$.

In our following explanation of the mechanism of the optimal regulation, we will set $\sigma = 1$ and let $\mu$ vary over time.

## 6.1 Capital Requirement for Banks' Indifference

In order to achieve an (interior; i.e., both types of banking sectors have a non-zero mass) equilibrium, a necessary condition is $\Pi(1, \frac{\lambda}{\lambda-1}; \mu, 1) = \Pi(r, s; \mu, 1)$ (it turns out that this interior solution is not always feasible; we will discuss this issue shortly). This constraint gives a mapping

$$\lambda = \Gamma(\mu) \tag{5}$$

If the condition (5) is satisfied, each bank is indifferent between the two types. If $\lambda > \Gamma(\mu)$, all banks prefer to be CB, and vice versa.

**Lemma 6.** *If $\frac{1-F(\hat{R})}{f(\hat{R})} > \frac{\int_{\hat{R}}^{\infty}(1-F(R))dR}{1-F(\hat{R})}$ (Condition 1) holds, $|MRS^{\Pi}|$ is decreasing in $\mu$.*

Condition 1 is not true for all distributions, but the next corollary states one sufficient condition.

**Corollary 2.** *If the hazard rate is increasing in $R$, $|MRS^{\Pi}|$ is decreasing in $\mu$.*

*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 (MsRP). As is well known, a necessary condition is that the hazard rate is decreasing in $\theta$: 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 Condition 1. □

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 3.** *If Condition 1 holds, the equilibrium regulation is pro-cyclical:* $\Gamma'(\mu) > 0$.

This result makes intuitive sense: As the risky asset return prospect improves (that is, $\mu$ increases), *ceteris paribus* it is more profitable for any given bank to be a shadow bank. So, to maintain the bank indifference, the government needs to make the CB sector more profitable. The government achieves that by loosening the capital requirement constraint – $\lambda = \Gamma(\mu)$, goes up. From the proof of Proposition 3, we can also see $r'(\mu) > 0$ and $s'(\mu) < 0$.

Figure 4 illustrates the procyclicality of equilibrium regulation. As the mean of the risky return $\mu$ goes up, the iso-profit curve of the banks $\Pi$ becomes flatter and the IR curve $s^u$ of the investors 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 $\lambda$ should go up.

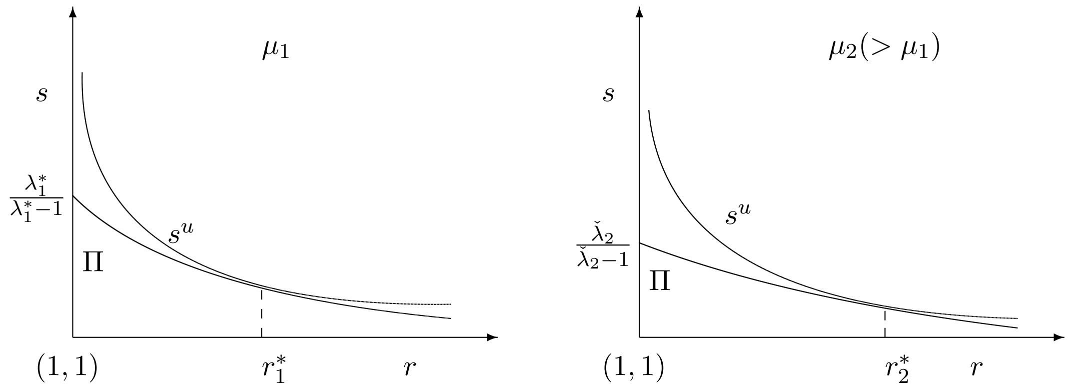


Figure 4: Procyclical Capital Regulation

## 6.2 Optimal Relative Banking Sector Size

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

**Lemma 7.** $l(\mu) > \Gamma(\mu)$

*Proof.* Condition (4) implies $r > 1$. Then since $\Pi_r < 0$ and $\Pi_s < 0$, the banks' indifference condition yields $\frac{\lambda}{\lambda-1} > s$ which implies $\lambda < \frac{s}{s-1} = l$. $\square$

The government seeks to achieve a target level of risk exposure of the economy based on the current conditions of the economy. In our model, the risk exposure of the economy is determined by the investment in the risky asset with respect to the total wealth of the economy, as in a conventional portfolio choice problem. To proceed, we define a state variable, the relative sector size $\xi \equiv \int_{SB} kdG / \int_{CB} kdG$. In our model, the fraction of the banking sector's net worth, $K/(W+K)$, where $K \equiv \int kdG$ is the total net worth of the banking sector, stays constant at its $t=0$ value $w$[12]. The investment in the risky asset is

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

Then, the fraction in the risky asset, $\alpha$, is

$$\alpha = \frac{w(l\xi + \lambda)}{\xi + 1}$$

This fraction specifies the realized risk exposure in any equilibrium (Note: $\alpha \in [w\lambda, wl]$, since $\xi \in [0, \infty)$). The optimal regulation equates the equilibrium frac-

[12]In an extension of this model to be discussed later, we let $w$ evolve over time.

tion with the target fraction. Denote the target risk exposure by $\phi(\mu)$, satisfying $\phi'(\mu) > 0$. By imposing $\alpha = \phi(\mu)$, we get

$$\xi = \frac{\phi(\mu) - w\Gamma(\mu)}{wl(\mu) - \phi(\mu)} \tag{6}$$

This equation equates the realized relative sector size to the target relative sector size. However, there is a caveat - we have to make sure that the right-hand side is non-negative. There are two cases in which it is negative. First, when the target risk exposure $\phi(\mu)$ is smaller than the lowest feasible realized risk exposure $w\Gamma(\mu)$ (note the feasible range of $\alpha$ above). In this case, even when all banks are CB (realized $\xi$ is zero), the realized risk exposure is higher than the target. Note, the government cannot set $\lambda$ lower than $\Gamma(\mu)$ because in that case all banks will flee to SB sector and the realized exposure will jump to $wl(\mu)$. The opposite case is that of $\phi(\mu) > wl(\mu)$. In this case, even when all banks are SB (realized $\xi$ is $\infty$), the realized risk exposure is lower than the target. We see that the mutual contracting between SB and their investors limit the maximum risk SB can assume. Henceforth, we will restrict our attention to the case where target risk exposure does not take such extreme values by assuming:

$$\phi(\mu) \in [w\Gamma(\mu), wl(\mu)]$$

First derivative of $\xi$

$$\xi'(\mu) = \frac{w(l(\mu) - \Gamma(\mu))}{(wl(\mu) - \phi(\mu))^2}\left[\phi'(\mu) - w\left(\frac{l'(\mu)\xi(\mu) + \Gamma'(\mu)}{\xi(\mu) + 1}\right)\right]$$

If $\phi'(\mu)$ is bounded away from zero for all values of $\mu$, there exists small enough $w$ such that $\xi'(\mu) > 0$. That is, the government wants a bigger SB sector as the risky asset has a better prospect[13].

## 6.3 Optimal Policy Schedule

The government aims to achieve the target $\xi$ of (6) without knowing $\mu$ directly, which is observed only by private sectors (note that $\xi$ is an observable to the government as well as the banks). This goal can be achieved by rearranging (6) to obtain a mapping $\mu = h(\xi)$ and then combine it with (5) to solve for $\lambda$

$$\lambda = \Gamma(h(\xi)) \equiv \Omega(\xi) \tag{7}$$

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

This is our optimal regulation specified in terms of the observable $\xi$. We now state the second main result of our paper.

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

*Proof.* $\Omega'(\xi) = \Gamma'(h(\xi))h'(\xi) = \frac{\Gamma'(h(\xi))}{\xi'(h(\xi))} > 0$ □

The regulation is stable in the following sense: Suppose the government announces the schedule $\lambda = \Omega(\cdot)$. Let us see what happens if $\mu$ goes up to $\mu'$. Then, $\lambda < \Gamma(\mu')$ so all commercial banks want to switch to SB. Although we do not model the process of switching their types, suppose some commercial banks have moved to the SB sector, implying in a higher $\xi$. In turn, it follows that $\Omega(\xi)$ and $\lambda$ also increase. It means that $\lambda$ gets closer to $\Gamma(\mu')$. This feedback process stops when $\lambda$ has risen to its new equilibrium value $\lambda' = \Gamma(\mu')$ with new mean and, by construction, we reach a new $\xi$, $\xi' = h^{-1}(\mu')$. Note the critical role of $\Omega'(\xi) > 0$ in transition to the new equilibrium.

This mechanism also has an implication for stabilizing the banking sector when the banking sector grows. For example, suppose that the risky asset gives a good return and accordingly the distribution $G$ moves to the right with no change in fundamentals. Then the SB sector grows relatively more than the CB sector because the former had more exposure to the risky asset. It follows that $\xi$ goes up and in turn $\lambda$ goes up by $\Omega(\cdot)$. This shift put a pressure on the SB sector so that some banks in the sector switch to the CB sector. This feedback lowers $\xi$ and $\lambda$, until $\xi$ and $\lambda$ return to their original level because the fundamentals did not change.

The same argument also shows that no bank coalition has any incentive to deviate in an equilibrium – Suppose a coalition of commercial banks arbitrarily decides to become unregulated. This increases the relative shadow banking sector size $\xi$. Following the schedule, capital requirement $\lambda$ increases. This makes the CB sector more profitable than the SB sector and consequently, the banks start to move to CB sector. This decreases $\xi$ and thus $\lambda$ until they reach the equilibrium values (values at the time the bank coalition decided to deviate).

# 7 An Example

We solve for equilibrium when the investors' preference is CRRA and risky asset return follows binomial distribution[14]:

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

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

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

The investors' participation constraint is

$$\mathbb{E}u\left(\min\{r, s\tilde{R}\}\right) = u(1). \tag{8}$$

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

$$\underline{R} < \frac{r}{s} < \overline{R} \tag{9}$$

Then, (8) yields the relation between $s$ and $r$

$$s^u(r) = \frac{1}{\underline{R}} \left( \frac{1-p}{1-pr^{1-\gamma}} \right)^{\frac{1}{\gamma-1}} \tag{10}$$

(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

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

[14]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 technicality without much gain.

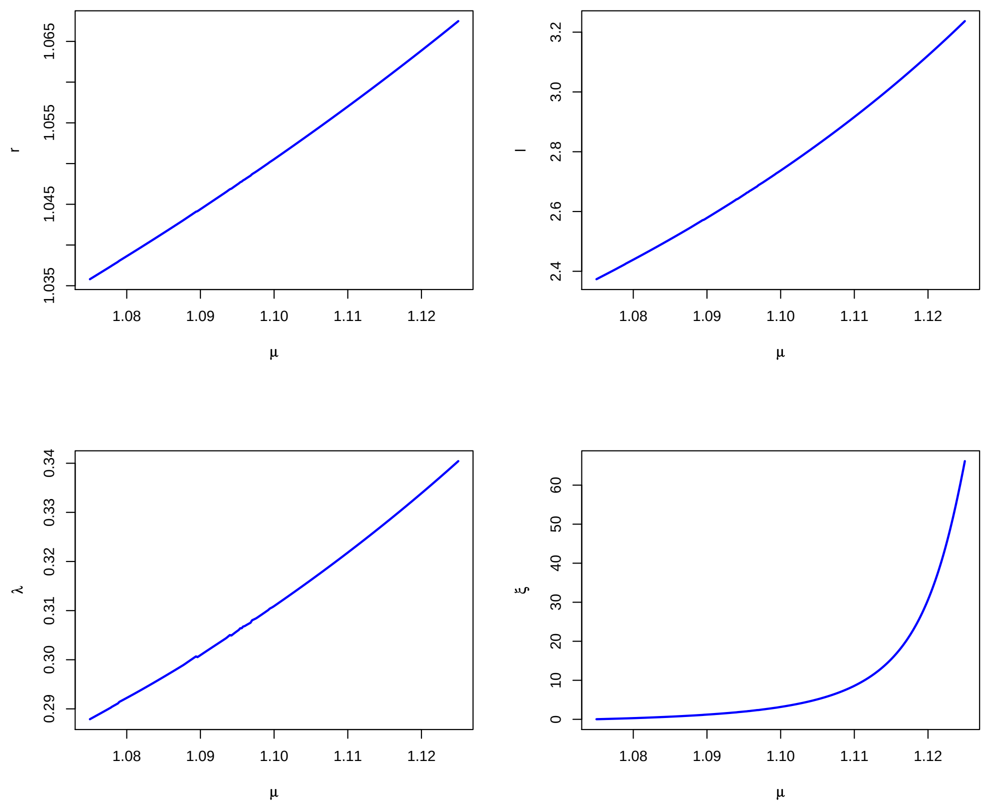


Figure 5: CRRA utility and Binomial returns: Comparative statics of equilibrium variables $r, l, \lambda, \xi$ ($\gamma = 4, \sigma = 0.15, \underline{R} = 0.4, w = 0.431$)

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

$$\left(\frac{r^{\gamma-1} - p}{1-p}\right)^{\gamma} = \left(\frac{r^{\gamma} - p\overline{R}}{\underline{R}(1-p)}\right)^{\gamma-1} \tag{11}$$

$$s = \frac{r^{\gamma} - pr}{r^{\gamma} - p\overline{R}} \tag{12}$$

Substituting $s$ from (10) into the left hand-side of (9), we get $1 < r$ and substituting $s$ from (12) into the right hand-side of (9), we get $r < \overline{R}$. So consistent with the assumption (9), we look for the solution(s) of (11) in the range $1 < r < \overline{R}$.

**Lemma 8.** *(11) has a unique solution in the range $1 < r < \overline{R}$.*

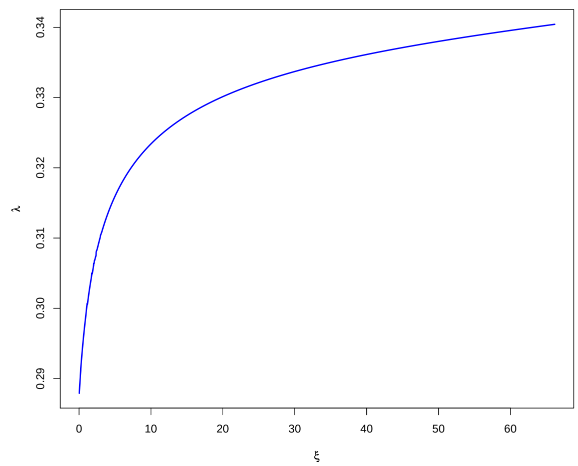


Figure 6: CRRA utility and Binomial returns: The Optimal Policy $\Omega$ ($\gamma = 4, \sigma = 0.15, \underline{R} = 0.4, w = 0.431$)

Equilibrium values of $\lambda$ and $\xi$ are given by

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

Next, we plot the various variables by fixing $\underline{R}$, varying $p$ and $\overline{R}$ such that $\mu \equiv p\overline{R} + (1-p)\underline{R}$ changes but $\sigma^2 \equiv p(1-p)(\overline{R} - \underline{R})^2$ stays constant. Accordingly, write $p(\mu) = \frac{(\mu - \underline{R})^2}{(\mu - \underline{R})^2 + \sigma^2}$ and $\overline{R}(\mu) = \mu + \frac{\sigma^2}{\mu - \underline{R}}$. Notice $\phi(\mu) = \frac{\mu - 1}{\gamma\sigma^2}$.

# 8 Robustness of the policy $\Omega$

Suppose due to a measurement error on the part of the government (either in the investor risk aversion coefficient $\gamma$ or in the return distribution $F$ of the risky asset), the government ends up implementing a policy which is *close* to the optimal policy. In the following, we show that the difference in risk exposure under this approximate policy and the optimal risk exposure is bounded.

Suppose there exists a sequence of policies $\{\Omega_n\}$ which is uniformly convergent to the right policy $\Omega$. 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 $\mu$, imagine the banking sector reaches the optimal $\xi^*$ if $\Omega$ was implemented. In contrast, if $\Omega_n$ satisfying the above condition is implemented, the banking sector will be settled at a suboptimal $\xi_n$ satisfying

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

If $\varepsilon$ is small enough, we can use a first-order approximation:

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

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 risk exposure of the economy is $\alpha(\xi) \equiv \frac{w(l\xi+\lambda)}{\xi+1}$. The values of $l$ and $\lambda$ are invariant for different policies in equilibrium, justifying $\alpha(\cdot)$ is only a function of $\xi$. The difference of risk exposures under the two policies is approximated using the above:

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

So, if $|\Omega'(\xi^*)|$ is bounded away from zero, the risk exposure is off the target in the order of $\varepsilon$. $\Omega'$ is bounded away from zero if $wl - \phi$ is bounded away from zero (see $\xi'(\mu)$ on page 21).

We can also ask what the expected social utility loss from this deviation is. The social expected utility $E\left[u\left(1 + \alpha(\tilde{R} - 1)\right)\right]$ is maximized at $\alpha = \alpha(\xi^*)$, and so $E\left[u'\left(1 + \alpha(\xi^*)(\tilde{R} - 1)\right)(\tilde{R} - 1)\right] = 0$. The EU loss is

$$E\left[u\left(1 + \alpha(\xi_n)(\tilde{R} - 1)\right) - u\left(1 + \alpha(\xi^*)(\tilde{R} - 1)\right)\right]$$
$$= (\alpha(\xi_n) - \alpha(\xi^*))\, E\left[u'\left(1 + \alpha(\xi^*)(\tilde{R} - 1)\right)(\tilde{R} - 1)\right] + O\left((\alpha(\xi_n) - \alpha(\xi^*))^2\right) = O(\varepsilon^2).$$

# 9 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 ($\gamma$).

We illustrate two approaches to hedge against this fluctuating risk aversion: with and without a detection technology of crisis. Neither of the approaches cannot 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 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 $\gamma$. 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.

## 9.1 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 $\gamma$ from its pre-crisis value $\gamma_0$[15]. We assume that even when the government is able to exogenously detect this jump, it cannot track $\gamma$ as it varies during the crisis before stabilizing to the end of crisis value $\gamma_1$ (potentially same as $\gamma_0$)[16]. 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, $\mu$, 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 $\gamma_0$). At time $t_1$, the following sequence of event unfolds: i) $\xi_{t_1}$ is realized and observed, ii) a crisis occurs iii) the government implements a new policy. Since the government observes $\xi_{t_1}$ before the crisis happens, it knows $\mu_{t_1}$. 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 $\xi$, limits the deviation from optimality.

Given $\mu$ is fixed, $l$ and $r$ in section 5 are functions of only $\gamma_t$, implying that (5) becomes

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

for some function $\Gamma^C$ where the superscript denotes a crisis. In the same manner, (6) is now written as

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

[15]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.

[16]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 $\xi$. If the government observes a sudden jump in $\xi$, it can suspect a shock to the economy which is not due to change in fundamentals $\mu, \sigma$. 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.

Following the same strategy in section 5, 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 $\gamma_{t_2}$ at the moment and implements the normal rule $\Omega_{\gamma_{t_2}}(\xi)$.

Certainly $\mu$ is also fluctuating during the crisis, 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.

## 9.2 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 $\mu$ is more stable than $\gamma$ in a crisis. It should be noted that $\gamma$ 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 from changing risk aversion, we can expand our set of policy instruments to $\xi_t$, $\xi_{t-1}$, $s_{t-1}$. It is very costly or even impossible for the government to observe the contemporaneous $\xi_t$ or $s_t$, but their past values, $\xi_{t-1}$ 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 $\mu_{t-1}$ and $\gamma_{t-1}$. 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 5, $\Omega(\xi_t; \gamma_t) = \Omega(\xi_t; \gamma_{t-1})$ in period $t$. Since $\gamma_{t-1}$ is a known function of $\xi_{t-1}$ 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.

# 10 Dynamic Environment

In this section, we consider a stylized dynamic setting of this model. By this extension, we show how the evolution of wealth of each agent influences our optimal regulation and what kind of modification is warranted. More importantly, the multi-period consideration enables us to incorporate the idea of traditional counter-cyclical regulation in an abstract form and clearly illustrate the sources of opposite policy suggestions.

The major difference we want to focus on in a dynamic setting is the time variation of the fraction of the banking sector's net worth, $w$, used in (6). In a static setting, we assumed that $w$ is a fixed small number. However, in a multi-period setting, the realized returns are varied across agents based on their terms of contracts, leading to a fluctuation in $w$. If $w$ changes over time, the optimal $\xi$ depends not only on $\mu$ but also on $w$ at the time. This dependence requires us to include $w$ as a state variable in our optimal regulation.

To demonstrate this effect, we make strong assumptions how agents behave in the new environment. The key assumption is that the world repeats itself at every period except only for the wealth levels of agents. Accordingly, given $\mu$, we have the same values of $\Gamma(\cdot)$, $\phi(\cdot)$, $s(\cdot)$, and $r(\cdot)$. We do not prove why this situation is optimal among all possibilities of contracts. As explained, the only part that is affected is the determination of function $h(\cdot)$ in (7), from $h(\xi)$ to $h(\xi, w)$. To guarantee $h$ is a well-defined function, we need the same condition:

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

because then the Jacobian determinant of $(\xi, w)'$ is always positive. Although this condition imposes the same restriction on $w$ (and $\phi(\cdot)$) as in the static case, here it leads to a restriction on asset returns, because $w$ depends on the history of realized returns. In a finite period case, at the minimum, we can claim that the partial derivative is positive if the return distribution has a corresponding upper bound.

The sequence of events at time $t$ is as follows:

1. $R_t$ is realized and observed.
2. $w_t$ is determined as a function of $R_t$ and $\mathcal{F}_{t-1}$, where $\mathcal{F}_{t-1}$ is the information set after all events at time $t-1$ take place.
3. $\mu_t$ is observed by market agents (not government).
4. Contract terms are determined based on economic conditions.
5. The banks choose their types under the regulation $\Omega(\xi, w_t)$.
6. A new equilibrium is achieved. This equilibrium delivers the equilibrium $\lambda_t$ and $\xi_t$, as before.

Before describing the optimal regulation in a dynamic case, we briefly explain the law of motion of $w$. If return $R_t$ is realized at time $t$ and $\sigma = 1$, the return (per capital) to each bank type is

$$R_t^{SB} \equiv \frac{1}{s_{t-1}-1}\max\{s_{t-1}R_t - r_{t-1}, 0\}$$
$$R_t^{CB} \equiv \max\{\lambda_{t-1}R_t + 1 - \lambda_{t-1}, 0\}$$

Then the 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 grow to

$$W_t = \frac{1}{s_{t-1}-1}K_{t-1}\frac{\xi_{t-1}}{\xi_{t-1}+1}(r_t^{SB}-1) + W_{t-1},$$

because the amount of wealth invested in SB, $\frac{1}{s_{t-1}-1}K_{t-1}\frac{\xi_{t-1}}{\xi_{t-1}+1}$, earns an excess return $r_t^{SB} \equiv \min\{r_{t-1}, s_{t-1}R_t\}$. Then

$$\begin{aligned} w_t &= \frac{K_t}{W_t + K_t} \\ &= \frac{R_t^{CB} + \xi_{t-1}R_t^{SB}}{\frac{1}{s_{t-1}-1}\xi_{t-1}(r_t^{SB}-1) + (\frac{1}{w_{t-1}}-1)(\xi_{t-1}+1) + (R_t^{CB} + \xi_{t-1}R_t^{SB})} \end{aligned}$$

As described above, we now have a new policy form, $\Omega(\xi, w)$. Even though $\xi$ and $w$ appear equivalent arguments in $\Omega(\cdot)$, it should be noted that their roles are entirely different. Since it is determined irrespective of the structure of the banking sector, $w_t$ plays no role in implementing stable and robust policies. It can be viewed that we have a menu of $\Omega(\xi)$ for different values of $w_t$ which is perfectly known before the banks make any decision at time $t$. In other words, the current wealth is a known value for investment decisions, while $\xi$ is a vehicle that leads to the economy to a fixed point which is designed to be optimal.

Another important aspect of this model is that the regulation is pro-cyclical. Cyclicality is purely determined by the sign of $\Gamma'$, which is independent of $w$. In our model, the fundamental force driving the cyclicality of regulation is the strategic behavior of banks between the two types of banks they can choose from. Hence, the cyclicality comes prior to considering implementing the optimality condition. 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 a counter-cyclical regulation policy based on the argument that the shadow value of bank capital rises in recessions. As they point out, this argument amounts to saying that the effect of capital crunches on the shadow value is bigger than that of deteriorated investment opportunities. Therefore, they conclude that a capital regulation should be loosened during recessions to alleviate the excessive scarcity of bank capital in bad times.

This standard idea in the last decade is embedded in our model. To see this, we shut down the shadow banking sector, $\xi = 0$ and let the government observe the business cycle ($\mu_t$). In this case, the government can achieve the optimal risk exposure by controlling $\lambda$ directly:

$$\lambda_t = \frac{\phi(\mu_t)}{w_t}.$$

Since $w_t$ is increasing in the realized return at time $t$, it is natural to assume that $\mu_t$ and $w_t$ are positively correlated. One rationale is that people update upward their belief on $\mu_t$ when they observe a positive shock to $R_t$. Kashyap and Stein (2004) also assume the same co-movement by describing recessions as having the lower stock of bank capital (low $w_t$) and fewer profitable lending opportunities (low $\mu_t$).

Now suppose that, in bad times, $w_t$ suffers a much bigger drop relative to $\phi(\mu_t)$. In other words, for a fixed $\lambda_t$, the banks lending ($\lambda_t w_t$) shrinks too much to achieve the desirable level of economic activity ($\phi(\mu_t)$), although $\phi(\mu_t)$ also goes down. This situation is the recessions described in Kashyap and Stein (2004). Then, in order to reach the optimal risk exposure, $\phi(\mu_t)$, the government has to increase $\lambda_t$, a looser regulation in bad times. This counter cyclical regulation mitigates a credit-freeze in bad times and a credit-craze in good times. More formally, we come to a counter-cyclical regulation if

$$\begin{aligned}
\frac{d\lambda_t}{dw_t} &= \frac{w_t \frac{d\phi(\mu_t)}{d\mu_t}\frac{d\mu_t}{dw_t} - \phi(\mu_t)}{w_t^2} \\
&= \frac{\frac{d\phi(\mu_t)/\phi(\mu_t)}{d\mu_t/\mu_t}\frac{d\mu_t/\mu_t}{dw_t/w_t} - 1}{w_t^2\phi(\mu_t)} < 0.
\end{aligned}$$

The inequality simply implies that $\phi(\mu_t)$ is more stable than $w_t$, a standard condition in the existing literature.

In our model, this advocacy of a counter-cyclical regulation can be valid only if there is no alternative type of banks – the unregulated bank. Once we open the conduit of capital toward the shadow banks, the effects in the literature disappear due to the competition between the two types of banks. Instead, with our optimal regulation, we showed $\frac{d\lambda_t}{d\mu_t} > 0$ or, equivalently, $\frac{d\lambda_t}{dw_t} > 0$, because $\mu_t$ and $w_t$ are positively correlated. 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.

# 11 Conclusion

In this paper, we model the decision of banks to become regulated. 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 achieves the optimal aggregate risk exposure 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 solution policy is shown to be pro-cyclical – loose capital requirement during good times and vice versa. This is opposite to the counter-cyclical policy proposed in literature. The reason for this dichotomy is that the literature ignores the ability of the banks to move in and out of the regulation umbrella, and our paper is a first step in this direction.

The proposed policy has several desirable features: (i) Macro-prudential – The objective of the policy is to control systemic risk, (ii) Market-based – Once the policy schedule is announced by the government, the market self-adjusts to attain the optimal risk exposure, (iii) Robust to business cycle fluctuations – Under the policy, economy transitions smoothly from one equilibrium to another as the underlying fundamentals change, and (iv) Robust to small 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 of this paper is the existence of regulatory arbitrage in the financial intermediation sector. 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 (we saw in section 5 how unless the desired risk exposure is too low, it is always optimal for the government to have both sectors exist in equilibrium). Moreover, insuring a huge commercial banking sector is 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 systemic risk.

We use a reduced-form setting that let us focus on the main forces driving our results. In particular, the two assets pay-off at $t = 1$ when the world ends and we abstract away from a complete description of the adjustment process of the economy as fundamentals change before $t = 1$. Risk-neutrality of banks and atomicity of investors help us simplify the respective portfolio optimization problems. Single-peaked risky return distribution simplifies the comparative statics of the terms of contract between the shadow banks and their investors. However, notice we do not appeal to any information asymmetry in this contract – the only financial friction is limited liability of the shadow banks. Even though we do not explicitly consider the cost bore by the government for providing the deposit insurance, note that our results hold for any government objective that delivers the form of setting the systemic risk exposure equal to an increasing function of mean $\mu$ of the risky return.

In a stylized multi-period extension of our model, we introduce one force that obtains the counter-cyclical 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 counter-cyclical if the shadow banking sector is turned-off, while it is pro-cyclical otherwise (that is, when the banks are allowed to be in the shadow). So, our pro-cyclical result holds over and above the traditional argument of counter-cyclical policy. Comparative statics with respect to the return variance $\sigma^2$ are ambiguous due to the following two opposing forces – On the one hand, the risk neutral banks dislike low $\sigma$ due to limited liability and on the other hand, the investors are willing to lend the banks more since they are safer when $\sigma$ is low. This limits our discussion of the optimal policy during crisis to spike in the investors risk aversion (as in 'flight to quality'). Lastly, some features of our policy rely on the sharp condition of banks' indifference under equilibrium.

# 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 Proposition 1** 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. 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 3**

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

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

$$\Pi(r,s) \geq \mu \qquad \forall s,r \tag{13}$$

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

$$\mu \geq r\int_{\hat{R}}^\infty dF \tag{14}$$

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

$$\begin{aligned}
\mu &\leq \frac{s}{s-1}\left[\int_{\hat{R}}^\infty (R-\hat{R})dF\right] \qquad \text{by (13)} \\
&= \frac{s}{s-1}\int_{\hat{R}}^\infty RdF - \frac{s}{s-1}\hat{R}\int_{\hat{R}}^\infty dF \\
&\leq \frac{s\mu}{s-1} - \frac{r}{s-1}\int_{\hat{R}}^\infty dF,
\end{aligned}$$

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 $r$.

$$\begin{aligned}\frac{\partial \Pi(r,s)}{\partial s} &= -\frac{1}{(s-1)^2}\left[\int_{\hat{R}}^{\infty}(R-\hat{R})dF\right] + \frac{r}{s(s-1)}\int_{\hat{R}}^{\infty} dF \\ &= -\frac{1}{s(s-1)}\Pi(r,s) + \frac{r}{s(s-1)}\int_{\hat{R}}^{\infty} dF \\ &\leq -\frac{\mu}{s(s-1)} + \frac{r}{s(s-1)}\int_{\hat{R}}^{\infty} dF \qquad \text{by (13)} \\ &\leq 0 \qquad \text{by (14)}\end{aligned}$$

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 ($MRS^{\Pi}$) is given by

$$MRS^{\Pi} = (s-1)\frac{1-F(\hat{R})}{\int_{\hat{R}}^{\infty}(R-r)dF}, \tag{15}$$

and positive since

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

as shown above.

The decreasing $MRS^{\Pi}$ 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**

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

$$J'(r) = \int_0^{\hat{R}} \left( \frac{u'(s^u R) s^u R}{u'(r) r} - 1 \right) dF - r \int_0^{\hat{R}} \frac{u'(s^u R) s^u R}{[u'(r) r]^2} [u''(r) r + u'(r)] dF$$

$$+ r \int_0^{\hat{R}} \frac{s^{u\prime}}{u'(r) r} [u''(s^u R) s^u R^2 + u'(s^u R) R] dF + 1$$

$$= \int_0^{\hat{R}} \left( \frac{u'(s^u R) s^u R}{u'(r) r} - 1 \right) dF - r \int_0^{\hat{R}} \frac{u'(s^u R) s^u R}{[u'(r) r]^2} [u''(r) r + u'(r)] dF$$

$$- (1 - F(\hat{R})) \frac{\int_0^{\hat{R}} [u''(s^u R) s^u R^2 + u'(s^u R) R] dF}{\int_0^{\hat{R}} u'(s^u R) R dF} + 1$$

$$= \int_0^{\hat{R}} \frac{u'(s^u R) s^u R}{u'(r) r} dF - F(\hat{R}) - r \int_0^{\hat{R}} \frac{u'(s^u R) s^u R}{[u'(r) r]^2} [u''(r) r + u'(r)] dF$$

$$- (1 - F(\hat{R})) \frac{\int_0^{\hat{R}} u''(s^u R) s^u R^2 dF}{\int_0^{\hat{R}} u'(s^u R) R dF} - (1 - F(\hat{R})) + 1$$

$$= \int_0^{\hat{R}} u'(s^u R) s^u R dF \left[ \frac{1}{u'(r) r} - \frac{u''(r)}{u'(r)} - \frac{1}{u'(r) r} \right] - (1 - F(\hat{R})) \frac{\int_0^{\hat{R}} u''(s^u R) s^u R^2 dF}{\int_0^{\hat{R}} u'(s^u R) R dF}$$

$$= - \frac{u''(r)}{u'(r)} \int_0^{\hat{R}} u'(s^u R) s^u R dF - (1 - F(\hat{R})) \frac{\int_0^{\hat{R}} u''(s^u R) s^u R^2 dF}{\int_0^{\hat{R}} u'(s^u R) R dF}$$

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

**Proof of Lemma 6** We rewrite (15) using a conditional expectation:

$$MRS^{\Pi} = - \frac{s - 1}{\frac{\int_{\hat{R}}^{\infty} R dF}{\int_{\hat{R}}^{\infty} dF} - r}$$
$$= - \frac{s - 1}{\mathrm{E}[R | R \geq \hat{R}] - r} \tag{16}$$

To see the effect of changing $\mu$, we use the parameterized distribution, $f^{\mu}$, as defined in the setup. Changing $\mu$ means horizontal translation, so $f^{\mu} = f(x - \mu)$. That is, all the distributions belong to the location family with $f$. We rewrite (16) with $\mu$ explicitly:

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

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

$$
\begin{aligned}
&-\int_{\hat{R}}^{\infty} Rf'(R-\mu)dR \cdot \int_{\hat{R}}^{\infty} f(R-\mu)dR + \int_{\hat{R}}^{\infty} Rf(R-\mu)dR \cdot \int_{\hat{R}}^{\infty} f'(R-\mu)dR \\
&= \hat{R}f(\hat{R}-\mu)\int_{\hat{R}}^{\infty} f(R-\mu)dR + \left[\int_{\hat{R}}^{\infty} f(R-\mu)dR\right]^2 - f(\hat{R}-\mu)\int_{\hat{R}}^{\infty} Rf(R-\mu)dR \\
&= \left[\int_{\hat{R}}^{\infty} f^{\mu}(R)dR\right]^2 - f^{\mu}(\hat{R})\int_{\hat{R}}^{\infty}(R-\hat{R})f^{\mu}(R)dR \\
&= \left[1-F^{\mu}(\hat{R})\right]^2 + f^{\mu}(\hat{R})(R-\hat{R})\left(1-F^{\mu}(R)\right)\Big|_{\hat{R}}^{\infty} - f^{\mu}(\hat{R})\int_{\hat{R}}^{\infty}\left(1-F^{\mu}(R)\right)dR \\
&= \left[1-F^{\mu}(\hat{R})\right]^2 - f^{\mu}(\hat{R})\int_{\hat{R}}^{\infty}\left(1-F^{\mu}(R)\right)dR, \qquad (17)
\end{aligned}
$$

where the first and third equalities come from the integration by parts. If (17) is positive, the iso-profit curve becomes flatter for higher $\mu$.

**Proof of Proposition 3** We have to consider two effects: changing $MRS^{\Pi}$ and shifting the investors' IR. First, consider the effect on $MRS^{\Pi}$ holding IR fixed. Lemma 6 shows that, if Condition 1 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 $\mu_1$. At a higher $\mu_2 > \mu_1$, the single crossing property implies that

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

where $\lambda_1^*$ denotes the equilibrium regulation at $\mu_1$ 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 $\mu$, 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 $\mu$. 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 $\mu$ and compute the total derivative of $s$ with respect to $\mu$.

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

Therefore, the partial derivative is

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

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 $\mu_2$ 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.

# References


[1] Acharya, Viral, 2009, A theory of systemic risk and design of prudential bank regulation *Journal of Financial Stability* 5, 224-255.

[2] Acharya, Viral, Philipp Schnabl, and Gustavo Suarez, 2010, Securitization without Risk Transfer, *NBER Working Paper Series, Working Paper 15730.*

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

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

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

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

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

[8] Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny, 2011, A Model of Shadow Banking, *NBER Working Paper Series, Working Paper 17115.*

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

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

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

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

[13] 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.

[14] Weitzman, Martin, 1974, Prices vs. Quantities, *Review of Economic Studies*, 41 (4), 477-491.