Dynamic Bank Capital Regulation in the Presence of Shadow Banks*

Arsenii Mishin†
HSE University

September 28, 2023

Abstract

Regulated banks expand relative to shadow banks in recessions and when credit spreads are high, while regulated banks shrink relative to shadow banks in expansions and when credit spreads are low. Motivated by these facts, I build a quantitative general equilibrium model with endogenous risk taking to study how competitive interactions between regulated banks and shadow banks affect optimal dynamic capital requirements. Limited liability and deposit insurance can lead regulated banks to provide socially inefficient risky loans when the returns on safer loans decline. Competition for scarcer funding can further lower the net returns on safe loans, making it more attractive for regulated banks to exploit the shield of limited liability with risky loans. Higher capital requirements can reduce inefficient risk at the cost of lower liquidity provision and some migration of credit from regulated banks to shadow banks. Accounting for the interactions of regulated and shadow banks can change the magnitude and direction of the optimal response of capital requirements to shocks that drive the business cycle. Moreover, Basel-III style rules that differentiate between the type of bank loans are much better at mimicking the Ramsey optimal capital requirements than standard rules that aggregate loans. The performance of such dynamic rules can be further improved once they are combined with a small static capital buffer.

JEL classification: E44, E58, G21, G23, G28, C54.

Keywords: Shadow Banks; Capital Requirements; Basel III; Liquidity Provision; Risk Taking.

*I am deeply indebted to Behzad Diba, Matthew Canzoneri, and Luca Guerrieri for their invaluable guidance, insightful discussions, comments, and patience. I also thank Mark Huggett, Minsu Chang, Dan Cao, Chad Curtis, and participants at seminars and conferences, including the macro advising group organized by Toshihiko Mukoyama, the Annual Conference of the Banco Central do Brasil, the Annual Canadian Economics Association Meetings, and the International Conference on Computing in Economics and Finance for helpful comments and discussions.

†Correspondence: Arsenii Mishin, Department of Theoretical Economics, HSE University, 11 Pokrovsky Bulvar, Moscow, Russia, 109028, phone: +7(495)772-95-90 (add 27774); email: aomishin@hse.ru

1 Introduction

The Basel III guidance proposes higher capital requirements in periods of excess aggregate credit growth to build up a greater cushion against potential losses in periods of stress.[1] These measures are aimed at increasing the resilience of the financial system to macroeconomic shocks. However, the financial system includes institutions that are not regulated as traditional banks. Shadow banks – lenders that perform bank-like functions but do not face capital requirements – have been expanding since the imposition of tighter bank regulatory standards in the wake of the 2007-09 global financial crisis.[2] This trend has raised concerns among policymakers about the effectiveness of capital regulation as loans can migrate from now more regulated traditional banks to unregulated shadow banks. How should prescriptions for optimal capital requirements be affected once competitive interactions of regulated and shadow banks are taken into account?

In this paper, based on data from the United States, I characterize three facts about the relation between regulated and shadow banks on business-cycle frequencies. I associate shadow banks with non-regulated intermediaries which provide liquidity services. I follow the methodology of Pozsar et al. (2013) to estimate the size of each of the banking sectors. First, shadow bank liabilities are significantly more volatile. Second, regulated banks expand relative to shadow banks in recessions, while regulated banks shrink relative to shadow banks in expansions. Third, higher credit spreads defined as average yields of corporate bonds relative to government bonds of similar maturity lead to expansion of regulated banks relative to shadow banks.

This paper develops and estimates a general equilibrium model with two types of financial intermediaries, which can account for all three facts. In the model, regulated and shadow banks borrow from households, issue equity and finance firms that undertake projects with varying degrees of risk. Households have preferences for liquidity offered by deposits of regulated and shadow banks. Each type of bank lends to two types of firms. Safe firms have a production technology that is only exposed to aggregate shocks; while risky firms are subject to both aggregate and idiosyncratic shocks. Idiosyncratic shocks have no influence on the average expected output of a risky firm but increase its variance. Regulated banks incur a cost to hide risky projects from regulators; while shadow banks do not have such penalty on risky returns. This feature that leads to a relatively higher social price of taking risk by

[1] See Basel Committee on Banking Supervision (2010).

[2] According to the recent data from the Financial Stability Board, assets of the narrow measure of shadow banking, which is estimated to pose the greatest financial stability risks, have increased by 75% globally since the global financial crisis. The United States constitutes the biggest share of the sector, amounting to 29% or \$15 trillion; this number is comparable to the size of assets of regulated banks in the country.

regulated banks reflects the availability of deposit insurance provided to regulated banks. The losses of their failure are borne by taxpayers, so risky projects are socially inefficient. Regulated banks take excessive risk by lending to risky firms. Unregulated shadow banks have no access to deposit insurance and face endogenously determined balance sheet constraints. Depending on the business cycle, banks find it privately optimal to either avoid idiosyncratic risk or load on this undiversified risk to exploit the shield of limited liability. Higher capital requirements can reduce or eliminate inefficient risk taking by forcing regulated banks to keep more "skin in the game" at the expense of lower liquidity provision and some migration of credit from regulated banks to shadow banks.

My contribution is threefold: First, I compute fully dynamic optimal capital requirements using a Ramsey approach. The Ramsey optimal capital requirements are set at the lowest level that prevents inefficient risk taking under the assumption that regulators can perfectly observe different types of exogenous shocks and commit to policy responses conditioning on shock realizations. I show that once we account for the interactions of regulated and shadow banks, optimal capital requirements react more aggressively to technology shocks. Moreover, introducing shadow banks changes the direction of the response of optimal capital requirements following capital quality shocks. In a model without shadow banks, optimal capital requirements rise in response to expansionary capital quality shocks, whereas accounting for shadow banks they contract.

Second, I provide an example in which an increase in capital requirements may lead to migration of credit from shadow to regulated banks as a result of a better designed capital regulation policy. It may occur when current capital requirements are set at suboptimally low levels, leading to socially inefficient lending. The optimal increase in capital requirements may drive the regulated bank deposit rate down due to higher convenience yield, decreasing the costs of providing loans and thus increasing lending of regulated banks.

Third, I explore how a bank regulator could implement optimal dynamic capital requirements in practice. I evaluate the so-called Basel III rule under which capital requirements respond to loan-to-output ratio. I find that a simple rule that differentiates between the two types of loans provided by each of the banking sectors does much better job at mimicking the response of the Ramsey optimal policy compared to a rule that conditions on aggregate loans. I consider both statistical measures, such as the R-squared, and welfare metrics, such as the consumption equivalent variation, that is computed by solving the model non-linearly. Although the rule that differentiates bank loans has a good fit with the Ramsey optimal policy, it leads to a number of consecutive inefficient risk-taking episodes. The reason is that the Ramsey policy completely switches off risk taking on the equilibrium path but the rule cannot always do that because of the informational requirements. This fact motivates me to study how a slightly elevated static capital requirement (capital buffers) can be combined with the considered simple rules. I find that combining the dynamic Basel-III style rules with capital buffers is welfare improving compared to using static buffers only, and the rule that aggregates loans now results in a smaller welfare loss.

Some general intuition regarding the model's main mechanisms can help shed light on all three contributions. Consider an example of a technology shock that reduces the returns on loans. There are two forces that are responsible for triggering a socially inefficient (or an excessive) risk-taking episode. One takes place independently of the presence of shadow banks, while another shows up due to the interactions of the two types of banks. First, lower loan returns decrease bank capital. Limited liability bounds possible further losses but does not affect potential gains. This asymmetry motivates bank managers to gamble for redemption by taking a flier on risky firms. In turn, deposit insurance acts as a subsidy for taking the risk because regulated banks do not internalize the probability of their default on the cost of borrowing. This combination of limited liability and deposit insurance allows regulated banks to pursue socially inefficient risk taking. Second, the model reflects the fact that regulated banks are less leveraged than shadow banks.[3] Accordingly, the returns on equity of regulated banks decline more modestly, inducing households to re-balance their equity portfolios. Households' equity holdings shift from shadow to regulated banks until the expected returns on the two types of equity are equalized. With binding capital requirements, this infusion of equity into regulated banks induces them to also attract more deposits to help fund an expansion of lending activity.[4] In the process, regulated banks drive up their deposit rate, which compresses the net returns on loans and thus boosts their risk-taking incentives. This interaction between bank and shadow banks magnifies the response of capital requirements.

To understand the reasons why accounting for the interaction between regulated and shadow banks could reverse the sign of optimal capital requirements, it is useful to separate the substitution and wealth effects of different shocks. For an expansionary productivity or capital quality shock, the wealth effect expands consumption, whereas the substitution effect expands investment. Depending on the balance between these two effects, the capital requirement could move in one direction or the other. Crucially, the substitution effect is stronger for technology shocks than for capital quality shocks regardless of whether shadow banks are in the model, but accounting for shadow banks increases the substitution effect.

[3] I follow Ferrante (2018) who estimates the leverage ratio of broker-dealers, expressed as assets to equity, to be between 20 and 40 compared to a value of 10 for regulated banks. Section 2 describes the shadow banking sector in more detail and further justifies this modeling feature.

[4] I derive formally that capital requirements always bind in the model. This result reflects households' preferences for bank debt, making debt a cheaper source of bank funding than equity.

In particular, the higher leverage of shadow banks implies that the returns on their equity are going to increase relatively more than the returns on regulated banks. This force allows them to attract additional funds from households reducing the financing costs for investment. For capital quality shocks, the substitution effect gets a boost sufficient to tilt the balance between the wealth and substitution effects. In response to an expansionary capital quality shock, shadow banks render the substitution effect so strong as to push up the price of installed capital, generating capital gains (the price of installed capital would have fallen in their absence). These capital gains also increase the returns on safe projects for regulated banks, and thus capital requirements can decline.

It may seem that a higher leverage of shadow banks drives the results. However, the interaction between regulated and shadow banks determines an additional transmission channel through household deposits. I estimate two key parameters that influence this transmission channel: the elasticity of substitution across the two types of deposits and the interest elasticity of total liquidity. A greater elasticity of substitution across the two types of deposits results in a smaller boost in capital requirements. In fact, regulated banks find it relatively cheaper to raise more debt with this greater elasticity of substitution because households require a smaller increase in the deposit rate for their switch away from shadow bank deposits. Moreover, a higher interest elasticity of total liquidity calls for a larger boost in capital requirements. In fact, when the supply of total liquidity is more elastic, a negative shock leads to a larger drop in total deposits. Since banks finance their loans from deposits, there is also a larger fall in loans. These greater contractionary effects result in a greater negative spillover of shadow banks on the net worth of regulated banks. The net worth of regulated banks decreases by more, leading to a greater response of capital requirements needed to combat the financing of inefficient loans.

This paper belongs to a growing literature on the design of optimal capital requirements in a quantitative general equilibrium framework (see, e.g., Van den Heuvel (2008); Collard et al. (2017); Mendicino et al. (2018); Davydiuk (2018); Begenau (2020); Malherbe (2020); Faria-e-Castro (2020); Canzoneri et al. (2020); Zhang (2020); Elenev et al. (2021); Begenau and Landvoigt (2021)). My main contribution is to evaluate the impact of the interactions between regulated and shadow banks on optimal capital requirements when banks endogenously choose the riskiness of their loan portfolios depending on the business cycle. Zhang (2020) also studies time-varying capital requirements in the presence of shadow banks. The author introduces a regulatory quadratic cost to model a capital requirement in an environment that features a costly state verification friction. I explicitly model moral hazard arising from limited liability and deposit insurance, incorporating it into a general equilibrium model with endogenous risk taking and aggregate uncertainty. I also contribute to the literature that studies Basel III prescriptions for setting the countercyclical capital buffer. My work stands out from this literature because I explicitly take into account the influence of shadow banks on devising Basel-III style rules.

The paper follows a tradition of the recent work on occasionally binding constraints as a source of non-linearity in financial crises such as Mendoza (2010), He and Krishnamurthy (2012), and Brunnermeier and Sannikov (2014). My model is in the vein of Boissay et al. (2016) and Gertler et al. (2020) that capture banking crises as rare events. I associate periods of excessive risk taking with regime shifts to model these rare episodes. In my framework, a low level of bank failures is linked to capital requirements being nearly optimal. Bank crises occur as a result of shocks that make socially inefficient lending attractive to banks, moving the economy into periods of excessive risk taking. These periods of inefficient lending are regime shifts. Optimal capital requirements respond to avoid these inefficient regime shifts.

To model the risk-taking incentives of the banking sector, I follow Van den Heuvel (2008) who shows how to exploit the shield of limited liability and deposit insurance to consider the financing choice, on the part of banks, of risky and safe projects. Van den Heuvel (2008) focuses on a setup that excludes aggregate risk. In this paper, I extend it to include aggregate risk, which enables me to study how risk-taking incentives vary over the business cycle. I model endogenously determined balance sheet constraints of shadow banks in the spirit of the work by Gertler and Karadi (2011), to which I add the possibility of default.

I model liquidity services provided by bank deposits through the utility function to capture the empirical evidence that deposits generate convenience yield (see, e.g., Gorton et al. (2012); Krishnamurthy and Vissing-Jorgensen (2012); Nagel (2016)).[5] Convenience yield is a non-pecuniary premium of holding safe and liquid assets that can perform money-like functions.[6] This modeling feature makes capital requirements always bind.[7]

To justify the difference in the returns from risky technologies across regulated and shadow banks, I relate to the literature that differentiates the expertise on the asset side of the two types of banks, such that regulated banks, protected by deposit insurance, have less incentives

[5]Feenstra (1986) shows that models with money-in-the-utility are functionally equivalent to models with transaction costs. Quadrini (2017) provides theoretical foundations for the bank liabilities channel that generates lower deposit interest rates paid by banks.

[6]In reality, shadow banks issue REPOs, commercial paper, MMMFs, and other short-term instruments in the money market to fund their assets. The recent empirical papers by Pozsar et al. (2013), Chernenko and Sunderam (2014), and Sunderam (2015) also find the importance of shadow banks for liquidity creation. To capture those benefits in the model, I put shadow bank deposits in the utility function. It also corresponds to the approach taken in Begenau and Landvoigt (2021).

[7]This paper belongs to the current stance of quantitative macro-banking research in general equilibrium which involves binding capital requirements. It is on a research agenda to construct quantitative models being consistent with reality that capital requirements do not bind all the time. For example, Begenau et al. (2020) develop a theory to capture several empirical facts, including unconstrained book and market leverage. At the same time, their model is partial equilibrium.

to develop a better screening system for selecting more creditworthy risky firms.[8]

It is common in the macro-banking literature to introduce a moral hazard problem into an environment that makes equity more expensive than debt due to either a tax distortion of bank profits or liquidity benefit of deposits. The moral hazard problem, which is associated with the presence of deposit insurance and limited liability, gives a non-negligible role for bank capital regulation. In this vein, Collard et al. (2017) examine jointly optimal prudential and monetary policies, which, unlike my setup, make excessive risk taking an off-equilibrium outcome. Davydiuk (2018) studies time-varying capital requirements within the Ramsey framework where banks take risks linked to the volume of credit. Here I consider how banks risk-shift by choosing projects of different quality. Begenau (2020) introduces the government subsidy of a particular functional form. One of the key elements that distinguish my work is that I explicitly derive the government subsidy associated with the provision of deposit insurance to banks and a share of non-defaulted shadow bank deposits. Canzoneri et al. (2020) compare simple implementable rules against the dynamic policy prescribed by a Ramsey planner in a model with endogenous risk taking. They find that a small static buffer performs as well as simple policy rules, including the prescription of Basel III. A common feature of those papers is that they abstract from shadow banks in their analyses.

Only few attempts have been made to differentiate between regulated and shadow banks in a general-equilibrium framework. The approach that the literature usually takes is to introduce the run-like behavior on shadow banks in the crisis time when banks are forced to liquidate assets at firesale prices. This environment amplifies the shocks and captures the highly non-linear nature of collapse. The papers that share these features include Gertler et al. (2016), Begenau and Landvoigt (2021), Ferrante (2018), Gertler et al. (2019), and Gertler et al. (2020). Begenau and Landvoigt (2021) find that a general equilibrium mechanism reduces the funding costs of banks following tighter capital requirements, and thus shadow banks expand their scale without becoming more risky as long as households care more about the overall liquidity than its composition. Unlike my work, they focus on static capital requirements and make no distinction between the technologies or expertise on the asset side of the two types of banks. Gertler et al. (2016) introduce a wholesale market where shadow banks borrow from retail banks and then characterize runs as self-fulfilling rollover crises in an infinite horizon endowment economy. Gertler et al. (2019) extend that framework to a conventional macroeconomic model, in which the banking sector is aggregated to include both investment banks and some large regulated banks. A pecuniary externality related to bank's

[8] For example, Buchak et al. (2018) show that shadow banks (represented by fintech lenders) serve more creditworthy borrowers. To mention more work that considers emergence of shadow banks based on differences in production technology of financial services, refer to Gertler et al. (2016), Ordoñez (2018), and Martinez-Miera and Repullo (2019), among others.

partial internalization of the effect of its leverage on the probability of a run for the whole banking system leads to a non-negligible role for capital regulation. The authors address the role of macroprudential policy in their companion paper Gertler et al. (2020) in which they relate credit growth to financial crises. Ferrante (2018) builds a model on Gertler et al. (2016) and introduces an informational friction that allows him to capture securitization and endogenous loan quality. He also considers the reintermediation of credit, but there is no direct link between his policies and capital requirements. Although I do not explicitly explore bank runs in my framework, I follow a complementary perspective and emphasize the role of shadow bank liquidity for capital requirements in a non-linear environment.

2 Stylized Facts

The definition of shadow banks has been evolving over the years and depends on the context. Before the global financial crisis, the term was meant to capture mainly those companies such as broker-dealers, mortgage finance firms, asset-backed regulated paper (ABCP) conduits, and money market mutual funds (MMMF) that participate in wholesale-funded, securitization-based lending process. Nowadays, many companies that used to be very different from banks have started getting involved in activities earlier associated with conventional banks. They include fintech lenders, insurance companies, private equity funds, hedge funds, and many others, all of which provide a significant source of credit to the economy. Since I focus on the liquidity provision function of financial institutions which society values in my framework, I measure the shadow banking sector by considering unregulated intermediaries the main function of which is to provide liquidity services.

I follow the methodology described by Pozsar et al. (2013) to estimate the size of regulated banks (RBs) and shadow banks (SBs) in the United States. I use quarterly data from the Flow of Funds for the period 1990 Q1 – 2022 Q1. This choice is governed by the fact that the shadow banking sector was relatively small before 1990, so its interactions with regulated banks were less relevant. Moreover, Basel I, which introduced capital requirements on banks, was issued in 1988 after the period of deregulation during 1980s.

To get a measure of shadow banks, I sum all the liabilities that relate to activities that are not backstopped by public guarantees such as asset-backed securities (ABS), repurchase agreements, security brokers and dealers, money market mutual funds, finance companies and commercial paper.[9] I do not consider GSE-backed mortgage pools because I would like to concentrate on the parts of shadow banks without explicit government sponsorship.[10] I

[9]See Appendix I.1 for more detail.

[10]Ferrante (2018) follows a similar approach.

use total liabilities of private depository institutions to estimate the size of regulated banks. I divide my measures by the implicit price deflator for GDP to convert the numbers into real terms.

I document three facts on the relation between the two types of banks and economic activity against which I will evaluate my modeling framework. Since I am interested in business-cycle frequencies, I measure shadow bank real liabilities, regulated bank real liabilities and real GDP as corresponding deviations from their HP trends.[11]

The first fact is that shadow bank liabilities are significantly more volatile; on average, they fall considerably more than regulated bank liabilities if output is below its trend, and they rise significantly more if output is above its trend. Table 1 summarizes the main results. It shows that the standard deviation of SB real liabilities is more than two times greater than the standard deviation of RB real liabilities when computed on business-cycle frequencies. Conditional on positive (negative) cyclical deviations of real GDP from its trend, SB real liabilities increase (decrease), on average, about nine times more than RB real liabilities increase (decrease).

Table 1: Summary Statistics

Statistics	Bank's Type: Real Liabilities	
	Regulated Banks	Shadow Banks
Standard Deviation	1.811 %	4.387%
Mean($\cdot$|GDP >0)	0.173%	1.544%
Mean($\cdot$|GDP <0)	-0.158%	-1.408%
Correlation with real GDP	-0.123	0.266

Note: All statistics are computed for HP-filtered shadow bank real liabilities, regulated bank real liabilities and real GDP measured as deviations from their HP trends. Mean($\cdot$|GDP >0) calculates the mean of bank real liabilities conditional on positive deviations of real GDP from its HP trend. Mean($\cdot$|GDP <0) calculates the mean of bank real liabilities conditional on negative deviations of real GDP from its HP trend.

The second fact links the measures of bank size with economic activity. Previous studies provide evidence that a considerable amount of debt migrated from the shadow banking system to the regulated banking sector during the Great Recession.[12] I show that this finding may not only be applied to particular episodes but it may also be generalized to the whole business cycle. To the best of my knowledge, this analysis has not been performed before in the literature.

[11] I select standard business cycle frequencies of 1,600 used for quarterly data. The results are also robust to the alternative choices of filters such as the bandpass filter.

[12] For example, He et al. (2010).

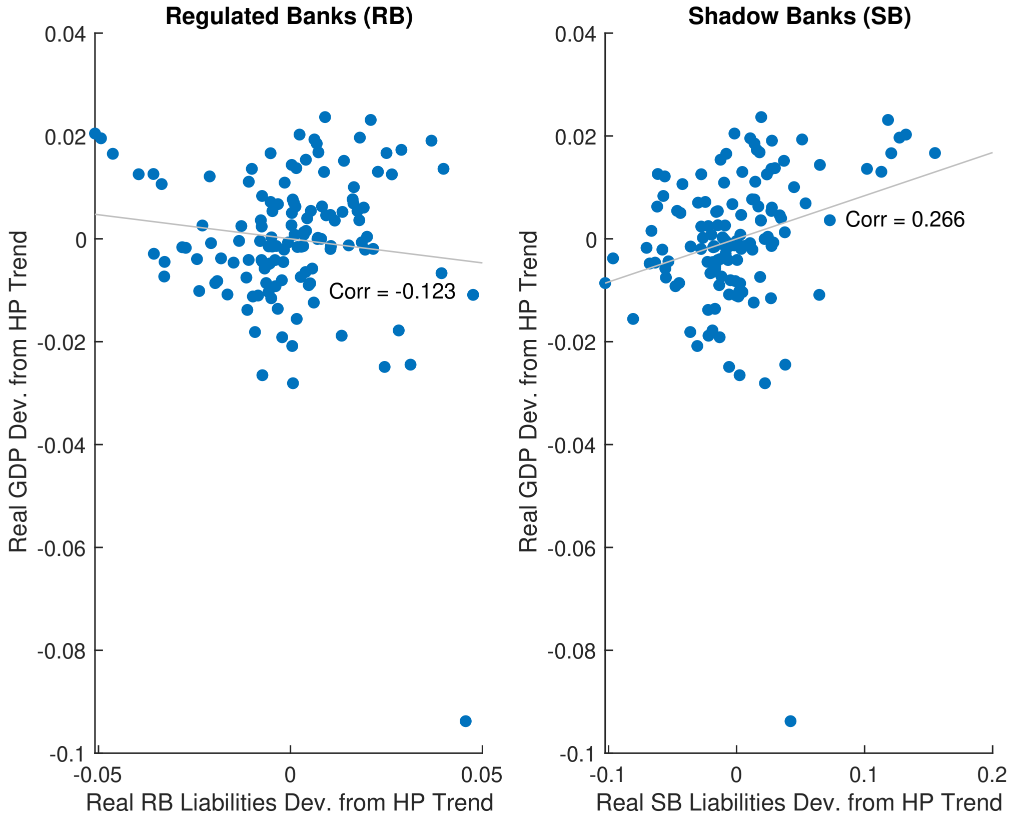

Figure 1. Bank Liabilities and Output

Note: The left (right) panel of the scatter plot shows the relationship between real GDP, expressed as a deviation from its HP trend, and regulated (shadow) bank real liabilities, expressed as a deviation from its HP trend.

Figure 1 shows that regulated bank liabilities are counter-cyclical, while shadow bank liabilities are pro-cyclical.[13] The size of RBs relative SBs falls with real GDP. As shown in Figure 2, the correlation between the difference in RB and SB real liabilities and real GDP is negative. RBs expand relative to SBs in recessions, while RBs shrink relative to SBs in expansions. I dub it the reintermediation channel that will be key to my risk-taking mechanism in the theoretical model.

The third fact highlights the relation between bank size and financial conditions. I link financial conditions to credit spreads defined as average yields of corporate bonds relative to

[13]It is interesting that the correlation coefficient of RB liabilities with GDP is negative because RBs, on average, expand when output increases and contract when output falls as indicated in Table 1. The negative correlation can be explained by the presence of large positive changes in RB liabilities when GDP is below its trend. Positive deviations of RB are more common but they are much smaller. They also concentrate around the center when GDP is above its trend. To lessen the influence of a choice of the filter on results, I repeat the procedure, applying the bandpass filter. The bandpass filter supports much stronger negative correlation of -0.3 for RB liabilities with GDP. This negative correlation is also robust to alternative choices of the sample period.

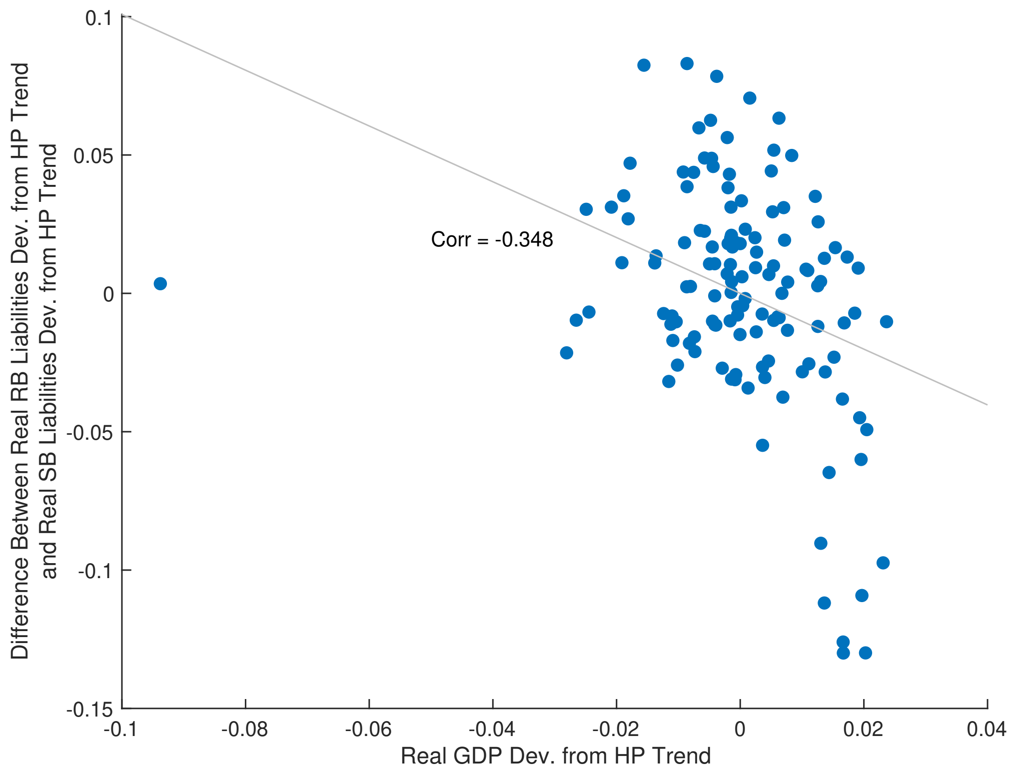

Figure 2. Output and Relative Size of Regulated Banks

Note: The scatter plot shows the empirical relationship between real GDP, expressed as a deviation from its HP trend, and the difference between RB and SB real liabilities, both measured as deviations from their HP trends.

government bonds of similar maturity. There are several measures proposed by the literature. As shown in the literature, high credit spreads are associated with times of elevated financial stress. I choose the monthly GZ credit spread index constructed by Gilchrist and Zakrajšek (2012). As demonstrated by the authors, it has the edge over widely used default-risk indicators such as the standard Baa–Aaa corporate bond credit spread in predicting future economic activity. I convert it to quarterly credit spreads. Figure 3 plots the dependence of RBs relative to SBs on the calculated credit spread.[14] This relationship is positive. The higher the spread (i.e., the tighter credit conditions), the more RBs expand relative to SBs.

[14] Since credit spreads are fast-moving variables, I measure real RB and SB liabilities as year-ahead deviations from their corresponding HP trends. This one-year shift is dictated by visually inspecting the time between the peaks and troughs of the credit spread and real RB and SB liabilities. A one-year ahead shift in real output is also used in Akinci and Queralto (2022) to relate real GDP to credit spreads. Since I do not use lags to link bank sizes with real output, I practically follow the same approach described in Akinci and Queralto (2022).

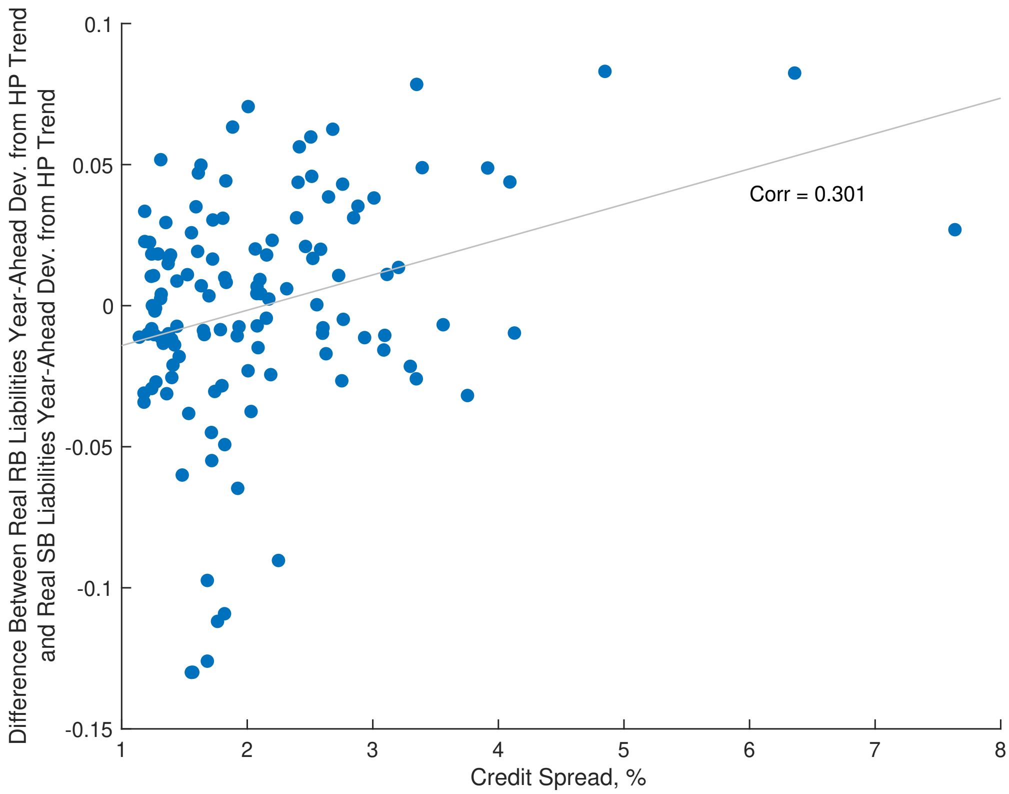

Figure 3. Credit Spread and Relative Size of Regulated Banks

Note: The scatter plot shows the empirical relationship between the credit spread, borrowed from Gilchrist and Zakrajšek (2012), and the year-ahead difference between RB and SB real liabilities, both measured as deviations from their HP trends.

3 Setup

The core framework is a standard RBC model extended to include regulated and unregulated banks that both have limited liability but only regulated banks enjoy deposit insurance. There are six types of agents in the model: households, shadow banks, regulated banks, production firms, capital producers, and a government that provides deposit insurance. To model shadow banks, I introduce the idiosyncratic shock and the possibility of default into the set up with the agency friction developed by Gertler and Karadi (2011). The description of regulated banks follows Canzoneri et al. (2020). As I show, the addition of shadow banks and consideration of their interactions with regulated banks have a substantial effect on model dynamics and policy implications.

3.1 Households

There is a continuum of infinitely lived households of mass one. Each household consumes, saves and supplies labor. Households consist of two types of members: workers and bankers. Each banker manages either a regulated bank or a shadow bank. Workers and shadow bankers can switch occupations each period. In particular, a shadow banker continues operating in the shadow banking sector next period with probability θ. The corresponding equal measure of workers randomly become shadow bankers. Upon exiting, shadow bankers return any earnings back to the household. There is perfect consumption insurance within the family. Households solve the following problem:

$$\max_{C_t, D_t^R, D_t^S, E_{s,t}^R, E_{r,t}^R, E_t^S} E \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma_c} - 1}{1-\sigma_c} + \sigma_0 \Psi\left(D_t^R, D_t^S\right) \right], \tag{1}$$

subject to

$$\begin{gathered} C_t + D_t^R + D_t^S + E_{s,t}^R + E_{r,t}^R + E_{H,t}^S = \\ W_t H + R_{t-1}^{dR} D_{t-1}^R + R_t^{dS} D_t^{\star S} + \Pi_t^S + R_{s,t}^{eR} E_{s,t-1}^R + R_{r,t}^{eR} E_{r,t-1}^R + R_t^{eS} E_{H,t-1}^S - T_t, \\ D_t^{\star S} = x_t^{\star} D_{t-1}^S, \\ E_{s,t}^R \geq 0, \\ E_{r,t}^R \geq 0. \end{gathered} \tag{2}$$

The representative household values consumption, C_t, and a mix of two types of deposits, D_t^R and D_t^S, where the superscript stands for the type of bank: a regulated (R) or a shadow (S) bank. Preferences for consumption are iso-elastic with $\sigma_c > 0$ governing the inverse of the intertemporal elasticity of substitution for consumption. The specification of derived utility from deposits is

$$\Psi\left(D_t^R, D_t^S\right) = \frac{\left(\alpha_1 \left(D_t^R\right)^{\sigma_d} + (1-\alpha_1)\left(D_t^S\right)^{\sigma_d}\right)^{\frac{1-\zeta}{\sigma_d}}}{1-\zeta}, \tag{3}$$

where $\alpha_1 \in [0,1]$ is the weight on regulated bank liquidity, $\frac{1}{1-\sigma_d} > 0$ is the elasticity of substitution between regulated and shadow bank debt, and $\zeta > 0$ is the measure of the elasticity of households supply of total deposits with respect to changes in the interest rate.

Workers supply labor, H, to firms inelastically for a wage, W_t. Households can acquire three types of bank equity, $E_{s,t}^R$, $E_{r,t}^R$, and $E_{H,t}^S$. R_t^{eS}, $R_{s,t}^{eR}$, and $R_{r,t}^{eR}$ with the corresponding gross returns R_{t+1}^{eS}, $R_{s,t+1}^{eR}$, and $R_{r,t+1}^{eR}$ next period. Banks differ by the riskiness of their investments.[15] Households invest into shadow bank equity and deposits through two funds,

[15] Banks can choose any level of risk (volatility) in a range from $\underline{\sigma}$ to $\bar{\sigma}$. I will show below that three types

which collect and distribute all the returns to the household family members. In particular, the equity fund supplies equal amount of equity to all existing shadow bankers. This feature ensures that both entering ("new") and existing ("old") shadow bankers start with the same ("shared") amount of equity, which makes them all ex-ante identical.[16] The deposit fund holds deposits in the shadow banks that it does not own. In the next period, the deposit fund receives the returns on non-defaulted deposits and distributes them equally to depositors. Regulated bank deposits are risk-free due to complete deposit insurance. In period t, they pay a non-contingent gross return R^{dR}_{t-1}. Shadow bank deposits pay a gross return R^{dS}_t on non-defaulted deposits $D^{\star S}_t$. The variable $x^{\star}_t$ denotes a non-defaulted share of shadow bank deposits and Π^S_t is the value that households receive after liquidating the assets of defaulted shadow banks. Households pay lump-sum taxes, T_t, collected by the government to provide deposit insurance.

Appendix H derives the first-order conditions of the household's problem. There are two features that make them different from the standard RBC equations. First, there is an additional benefit from holding deposits. This fact becomes relevant in what follows. Second, the state-contingent required rates of return on the three types of equity play an important role for allocation of capital across firms' technologies in equilibrium.

3.2 Banking Sector

There is a continuum of measure one of each type of bank, $i = R, S$, indexed by $j \in [0, 1]$. Banks lend to a mix of two types of production firms. One type is subject to aggregate shocks only (safe firms, for short). Safe loans earn R^l_{t+1} next period. Another type is subject to both aggregate and idiosyncratic shocks (risky firms, for short). Banks finance these loans by raising deposits and equity from households. Risky loans yield $R^l_{t+1} + \frac{\varepsilon^i_{j,t+1}}{Q_t}$ where Q_t is the price of capital and $\varepsilon^i_{j,t+1}$ is the idiosyncratic shock that has no influence on average expected output but increases its variance. Equipped with limited liability, banks have incentives to increase exposure to the idiosyncratic shocks. Regulated banks lose, on average, ξ_R from financing risky firms because it is relatively costly for them to hide risky projects from the regulator. Moreover, each type of bank has its own variance of the idiosyncratic process, reflecting the difference in available technologies. A banker j of type i creates a loan portfolio with riskiness $\sigma^i_{j,t}$ and earns total returns $R^l_{t+1} + \sigma^i_{j,t}\frac{\varepsilon^i_{j,t+1}}{Q_t}$ by directing a fraction $\sigma^i_{j,t}$ of loans

of equity span all possible equilibrium risk choices for banks.

[16]It goes beyond the scope of this paper to consider a heterogeneous-agent framework in which the distribution of shadow bank equity holdings also plays a role.

to a risky firm and the remaining share $1-\sigma^{i}_{j,t}$ to a safe firm.[17,18]

3.2.1 Shadow Banks

In period t, a shadow bank j receives $e^{S}_{j,t}$ units of equity from the equity fund, demands $d^{S}_{j,t+1}$ units of deposits from the deposit fund and lends $l^{S}_{j,t}$ to firms. At time $t+1$, the bank's net cash flow, $\omega^{S}_{j,t+1}$, is the difference between earnings on loans and payments on deposits:

$$\omega^{S}_{j,t+1} = \max\left[\left(R^{l}_{t+1} + \sigma^{S}_{j,t}\frac{\varepsilon^{S}_{j,t+1}}{Q_t}\right) l^{S}_{j,t} - R^{dS}_{t+1} d^{S}_{j,t+1}, 0\right]. \tag{4}$$

If the net cash flow is positive, the bank pays it out to households in dividends. If the net cash flow is negative, limited liability protects the bank from making any losses, so the bank gets zero and liquidates its assets to partially reimburse depositors.

Since households have preferences for the bank debt, shadow banks find it cheaper to finance their assets by issuing deposits only. To motivate shadow banks to accumulate equity, I consider a costly enforcement problem in the spirit of Gertler and Karadi (2011). In particular, at the start of each period, the shadow banker can choose to transfer the share λ of loans back to the household. If the banker opts for the transfer, the deposit fund that represents the interests of the depositors of another household can force the banker to liquidate the remaining fraction $1-\lambda$ of the assets. In effect, depositors provide funds to the shadow bank if the following incentive-compatibility constraint holds:

$$V^{S}_{j,t} \geq \lambda l^{S}_{j,t}, \tag{5}$$

where $V^{S}_{j,t}$ is the value of the bank measured by the expected terminal wealth.

Let $\hat{E}_t$ denote expectation taken over the joint distribution of $\varepsilon^{R}_{j,t+1}$ and R^{l}_{t+1}, subject to information known at time t. The objective of the bank is to maximize its value:

$$V^{S}_{j,t} = \max_{l^{S}_{j,t+i},\sigma^{S}_{j,t+i}} \hat{E}_t \left\{ \sum_{i=0}^{\infty} (1-\theta)\,\theta^{i} \Lambda_{t,t+1+i} \omega^{S}_{j,t+1+i} \right\}, \tag{6}$$

[17]The firm section describes the production functions of each type of firm, from which I derive the returns to loans and show how they compose the portfolio returns postulated here.

[18]The statement that a regulated bank only deals with one safe and one risky firm comes at no loss of generality since diversification is useless given constant returns to scale technology of safe firms and detrimental for loans to risky firms. You can consult Collard et al. (2017) for a more formal exposition.

subject to

$$
\begin{aligned}
V_{j,t}^{S} &\geq \lambda l_{j,t}^{S}, \\
\omega_{j,t+1+i}^{S} &= \max\left[\left(R_{t+1+i}^{l} + \sigma_{j,t+i}^{S} \frac{\varepsilon_{j,t+1+i}^{S}}{Q_{t+i}} - R_{t+1+i}^{dS}\right) l_{j,t+i}^{S} + R_{t+1+i}^{dS} e_{t+i}^{S}, 0\right], \qquad (7) \\
\underline{\sigma}^{S} &\leq \sigma_{j,t} \leq \bar{\sigma}^{S},
\end{aligned}
$$

where $\Lambda_{t,t+i} = \beta^{i} \frac{\lambda_{ct+i}}{\lambda_{ct}}$ is the stochastic discount factor applied to earnings at $t+i$.

I make the following assumptions:

Assumptions (Shadow Banks):

1. At least a positive fraction $\underline{\sigma}^{S}$ of the value of total loans will go to risky firms and at least a positive fraction $1-\bar{\sigma}^{S}$ of the value of total loans will go to safe firms.
2. $\varepsilon_{j,t+1}^{S}$ follows a Normal distribution with mean zero and variance τ_{S}^{2}.
3. The amount of equity that the shadow banker receives from the equity fund is the same across all shadow banks, i.e. $e_{j,t}^{S} = E_{t}^{S}$ for all $j \in [0,1]$.
4. Shadow banker exits with i.i.d. probability $1-\theta$ next period.

The first assumption takes the form of the minimum scale condition for the allocation of capital across two production functions. It ensures that the returns to both safe and risky loans are always defined. The second assumption makes it possible to represent the expectation operator $\hat{E}_{t}$ as an expectation taken over the aggregate shock, nesting expectations taken with respect to the idiosyncratic shock. The unbounded support of the distribution does not require any restrictions on the parameter values when measuring the impact of aggregate shocks. The third assumption allows me to model the shadow banking sector within a representative agent framework. The final assumption is used to limit the ability of shadow banks to get rid of the financial friction.

3.2.2 Regulated Banks

A regulated banker j enters period $t+1$ with $l_{j,t}^{R}$ units of loans, financed by issuing $d_{j,t}^{R}$ units of deposits and $e_{j,t}^{R}$ units of equity. The balance sheet condition is

$$
l_{j,t}^{R} = e_{j,t}^{R} + d_{j,t}^{R}. \qquad (8)
$$

The net worth, $e^R_{j,t}$, available to banks at the end of period t (going into period $t+1$) evolves according to:

$$e^R_{j,t+1} = \max\left[\left(R^l_{t+1} + \sigma^R_{j,t}\frac{\varepsilon^R_{j,t+1}}{Q_t} - R^{dR}_t\right) l^R_{j,t} + R^{dR}_t e^R_{j,t}, 0\right] - z_{j,t+1} \tag{9}$$

where $z_{j,t+1}$ is the net payout to the bank's shareholders in $t+1$ after the realization of the shocks in t.

The banker's objective is to maximize the expected discounted sum of equity payouts:

$$V^R_{j,t} = \max_{l^R_{j,t+i}, e^R_{j,t+i}, \sigma^R_{j,t+i}} \hat{E}_t \left\{ z_{j,t} + \sum_{i=0}^{\infty} \Lambda_{t,t+1+i} z_{j,t+1+i} \right\}, \tag{10}$$

subject to

$$\begin{aligned}
e^R_{j,t+i} &\geq \gamma_{t+i} l^R_{j,t+i}, \\
z_{j,t+1+i} &= \max\left[\left(R^l_{t+1+i} + \sigma^R_{j,t+i}\frac{\varepsilon^R_{j,t+1+i}}{Q_{t+i}} - R^{dR}_{t+i}\right) l^R_{j,t+i} + R^{dR}_{t+i} e^R_{j,t+i}, 0\right] - e^R_{j,t+i}, \\
l^R_{j,t+i} &\geq 0, \\
\underline{\sigma}^R &\leq \sigma^R_{j,t+i} \leq \bar{\sigma}^R.
\end{aligned} \tag{11}$$

The capital requirement stipulates that equity needs be at least the fraction γ_t of total loans for the bank to operate in each period. The non-negativity constraint on the amount of loans excludes the possibilities of short-selling.[19]

I make the following assumptions:

Assumptions (Regulated Banks):

1. At least a positive fraction $\underline{\sigma}^R$ of the value of total loans will go to risky firms.
2. The bank supervisory authority will prevent risky loans in excess of financing a share $\bar{\sigma}^R$ of total loans where $\underline{\sigma}^R < \bar{\sigma}^R < 1$.
3. The CDF of $\varepsilon^R_{j,t+1}$ denoted by G is Normal with mean $-\xi_R < 0$ and variance τ^2_R.

The first assumption is needed for the same reasons as discussed in the shadow banking section. The second assumption states that the regulator observes excessive risk taking imperfectly. The threshold for the share of risky loans from which the authority starts detecting excessive risk is described by $\bar{\sigma}^R$. The interpretation is that the regulator has power

[19] The reasons for no short-selling constraint come from the fact that bank's objective function is convex in risk. Section 4 clarifies this point.

to penalize banks significantly enough to make them never find it optimal to finance the fraction of risky firms greater than $\bar{\sigma}^R$. The third assumption formalizes inefficient risk taking by making the expected return on a loan portfolio decrease in risk.

3.3 Production Firms

Competitive firms are owned by households and produce goods using capital and labor as inputs. There are two classes of firms, safe and risky, each having measure one. Firms borrow from banks to purchase capital. Next period they collect income from production activity and the sale of undepreciated capital. They distribute the resulting payoff to workers and the banks that they serve. Banks perfectly observe firms' output and realizations of the idiosyncratic shock, so they can enforce the full payment of firms' payoffs.

The production function of safe firms is

$$Y^s_{j,t} = A_t K^{\alpha}_{j,t} H^{1-\alpha}_{j,t}, \tag{12}$$

where A_t is the aggregate technology shock.

The production function of risky firms is

$$Y^r_{j,t} = A_t K^{\alpha}_{j,t} H^{1-\alpha}_{j,t} + \varepsilon_{j,t} K_{j,t}, \tag{13}$$

where $\varepsilon_{j,t}$ is the idiosyncratic shock specific to firm j. It corresponds to the idiosyncratic shock that was introduced in Section 3.2.

Let $\pi_{j,t+1}$ denote the revenue of firm j in period $t+1$ net of expenses.

$$\pi_{j,t+1} = Y_{j,t+1} + (1-\delta) Q_t K_{j,t+1} - W_{t+1} H_{j,t+1} - R^l_{j,t+1} l_{j,t}. \tag{14}$$

The term $Y_{j,t+1}$ is output in period $t+1$, Q_t is the price of capital in terms of the final good, δ is the depreciation rate, $H_{j,t+1}$ is the labor input in production, W_{t+1} is compensation for labor, and $R^l_{j,t+1}$ is the borrowing rate. Firms maximize expected profits, knowing that they will be able to choose the optimal quantity of labor $H_{j,t+1}$ next period:

$$\max_{l_{t+i},\, K_{j,t+i+1}} E_t \left\{ \sum_{i=0}^{\infty} \Lambda_{t,t+i} \max_{H_{j,t+i+1}} \pi_{j,t+i+1} \right\}, \tag{15}$$

subject to

$$Q_{t+i} K_{j,t+i+1} = l_{j,t+i}. \tag{16}$$

Appendix F derives the first-order conditions and shows that individual firms can be aggregated into representative firms with the same capital-to-labor ratio. It finds that

$$R_t^l \equiv R_{j,t}^l = \frac{\alpha A_t}{Q_{t-1}}\left(\frac{K_t}{H_t}\right)^{\alpha-1} + (1-\delta)\frac{Q_t}{Q_{t-1}}, \tag{17}$$

$$R_{j,t}^{rl} = R_t^l + \frac{\varepsilon_{j,t}}{Q_{t-1}}. \tag{18}$$

Equation (17) defines the returns on safe loans composing the rental rate on capital and capital gains from re-selling undepreciated capital. Equation (18) expresses the returns on risky loans as the sum of the returns on safe loans and the extra return/loss from the realization of the idiosyncratic shock normalized by the price of capital.

3.4 Capital Producing Firms

At the beginning of period t, after realization of the shocks, competitive capital producing firms buy capital from production firms, repair depreciated capital and build new capital. They sell both the new and re-furbished capital at the end of period t. The cost of replacing worn out capital is unity. There is a common market for capital for safe and risky firms. The value of a unit of new capital is Q_t.

Let I_t^g denote aggregate gross investment expenditures. There are quadratic adjustment costs measured in units of investment. Aggregate investment expenditures of size I_t^g yield net investment of size I_t^n such that

$$I_t^n = \left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] I_t^g, \tag{19}$$

where ϕ is a parameter that governs adjustment costs. The functional form is based on Christiano et al. (2005). The aggregate capital stock evolves according to:

$$K_{t+1} = I_t^n + (1-\delta)K_t, \tag{20}$$

where K_{t+1} is the total capital allocated to the representative firms.

Capital producing firms solve:

$$\max_{I_{t+i}^g} E_t \sum_{i=0}^{\infty} \Lambda_{t,t+i}\left[Q_{t+i}\left[1 - \frac{\phi}{2}\left(\frac{I_{t+i}^g}{I_{t+i-1}^g} - 1\right)^2\right] I_{t+i}^g - I_{t+i}^g\right], \tag{21}$$

where Q_t is given and $\Lambda_{t,t+i} = \beta^i \frac{\lambda_{ct+i}}{\lambda_{ct}}$ is the stochastic discount factor of households.

3.5 The Government

Deposit insurance requires the government to raise taxes. Given the Ricardian nature of the model, positing the availability of lump sum taxes T_t implies that the government budget can be balanced period by period without loss of generality. Appendix G shows the equilibrium tax necessary to support the insurance scheme.

4 Analytical Results

In this section, I derive some analytical results that will be used to solve the model numerically. I apply the results from the Appendix to characterize the problem of each type of bank and the model's equilibrium. I also discuss the mechanism that generates socially inefficient risk taking which can be curbed by capital requirements.

4.1 Shadow Banks: Characterization

Appendix D finds that the objective function is increasing in σ_t^S, so each shadow banker opts for financing the maximum share of risky projects. It also establishes that there is a representative shadow banker.

The value of the representative shadow bank can be separated into a component that depends on loans, υ_t, and a component that depends on net worth, η_t, shown in the Appendix:

$$V_t^S = \upsilon_t l_t^S + \eta_t e_t^S. \tag{22}$$

The financial friction endogenously restricts lending by the amount of equity. Even though shadow banks do not choose equity on their own, they still have incentives to accumulate equity in order to limit the effect of the financial friction. These incentives are shared with the equity fund that represents their interests perfectly. In fact, the equity fund redistributes existing equity and receives additional equity that is optimally supplied by the household.

To make the financial friction relevant, consider $0 < \upsilon_t < \lambda$ where the bank equates the benefit from diverting funds with its cost and thus the incentive constraint binds. Then

$$l_t^S = \frac{\eta_t}{\lambda - \upsilon_t} e_t^S = \phi_t^S e_t^S, \tag{23}$$

where $\phi_t^S = \frac{l_t^S}{e_t^S}$ is the leverage of the banker.

4.2 Regulated Banks: Characterization

Appendix E.1 derives the first-order conditions of regulated banks on their choices of loans, deposits and equity. Appendix E.3 proves that

Proposition 1. *In equilibrium, capital requirements always bind, i.e.* $e^R_{j,t} = \gamma_t l^R_{j,t}$.

This proposition formalizes the argument that additional benefit from holding deposits makes debt a cheaper source of bank funding than equity. So regulated banks will issue as much deposits as is allowed by the capital regulation.

4.3 Equilibrium Characterization

I consider a competitive equilibrium in which agents take aggregate prices as given. Appendix E.2 proves that

Proposition 2. *A regulated bank j optimally considers only two values for its choice of risk, i.e.* $\sigma^R_{j,t} = \underline{\sigma}^R$ *or* $\sigma^R_{j,t} = \bar{\sigma}^R$ *in equilibrium.*

Note that this proposition does not state that each bank chooses the same amount of risk. It is possible that some positive fraction of regulated banks finance the maximum share of risky projects, while the remaining fraction of regulated banks choose the minimum share of risky projects in equilibrium.

Formally, let μ_t be a share of regulated banks that finance the fraction $\bar{\sigma}$ of risky projects. Following the notation from the households' section, μ_t also corresponds to the share of household's equity invested into risky banks out of the total equity allocated to regulated banks, i.e.

$$\mu_t = \frac{E^R_{r,t}}{E^R_{r,t} + E^R_{s,t}}. \tag{24}$$

Three types of equilibria are possible:

1. Safe equilibrium: $\mu_t = 0$.

2. Risky equilibrium: $\mu_t = 1$.

3. Mixed equilibrium: $0 < \mu_t < 1$.

If short-selling were allowed, and regulated banks were in the safe equilibrium, risky loans would be overpriced over safe loans because expected returns on risky loans are lower in the safe equilibrium. Hence, each bank would want to short sell risky loans (which means that it would acquire a negative amount of risky loans), leading to an arbitrage opportunity. A similar reasoning applies to the risky equilibrium, in which shorting safe loans would result in arbitrage profits. Thus, the condition that loans cannot be negative is needed to exclude any arbitrage opportunities. In the mixed equilibrium, the expected dividends of risky and safe banks are equal, so a positive measure of regulated banks (call them risky banks) choose to finance the maximum share, $\bar{\sigma}^R$, of risky projects and a positive measure of regulated bankers (call them safe banks) opt for financing the minimum share, $\underline{\sigma}^R$, of risky projects.

Since each bank within a group (safe or risky) is alike and receives an aliquot share of financing, the bank-specific terms and aggregate terms are related as follows:

$$E_{r,t}^R = \int_0^{\mu_t} e_{j,t}^R dj, \tag{25}$$

$$E_{s,t}^R = \int_{\mu_t}^{1} e_{j,t}^R dj. \tag{26}$$

Appendix E.4 establishes that the value function of regulated banks is linear in loans. It implies that there is a representative regulated bank of type $i = s, r$ that finances the fraction $\underline{\sigma}^R$ of risky projects when $i = s$ and finances the share $\bar{\sigma}^R$ of risky projects when $i = r$.

4.4 Discussion of the Excessive Risk-Taking Mechanism

There are two forces that trigger an excessive risk-taking episode. The first one comes from the direct effects of shocks on a loan portfolio of regulated banks. The second force is a result of the interactions of regulated and shadow banks. I will discuss each of them in turn.

The Appendix decomposes the expected dividends of regulated banks into two components:

$$\Omega\left(\mu_t, \sigma_t^R;\, l_t^R\right) = E_t\left\{\Lambda_{t,t+1} l_t^R \left[\omega_1 + \omega_2 - (1-\gamma_t)\right]\right\}, \tag{27}$$

where I omit the index of regulated bank and

$$\omega_1 = \left(R_{t+1}^l - R_t^{dR}(1-\gamma_t) - \xi_R \sigma_t^R\right)\left(1 - G(\varepsilon_{t+1}^*)\right), \tag{28}$$

$$\omega_2 = \sigma_t^R \frac{\tau_R}{\sqrt{2\pi}} e^{-\left(\frac{\varepsilon_{t+1}^* + \xi_R}{\tau_R\sqrt{2}}\right)^2}. \tag{29}$$

The cutoff point ε_{t+1}^* below which the bank's net worth is negative is defined by $R_t^{dR}(1-\gamma_t)\,Q_t - R_{t+1}^l Q_t = \sigma_t^R \varepsilon_{t+1}^*$.

The first component, ω_1, is the expected net income on loans: the bank receives $R_{t+1}^l - R_t^{dR}(1-\gamma_t) - \xi_R \sigma_t^R$ if it does not use the deposit insurance which happens with probability

$1 - G(\varepsilon^*_{t+1})$; it gets zero otherwise. The term $\xi_R \sigma^R_t$ is the average reduction of the returns for financing risky firms. The second component, ω_2, accounts for the benefits from limited liability. It captures the additional effect of volatility on the risky returns. The bank views projects as a call option, and thus ω_2 is increasing in τ_R.

Risk-taking incentives depend on the net returns on loans. When they decrease, regulated banks find it more attractive to use limited liability that shields them from downside risk. So banks are tempted to take a flier on risky loans.

How do shadow banks affect the risk-taking incentives of regulated banks? Shadow banks are more leveraged, and thus their net worth is more negatively affected than the net worth of regulated banks after a negative shock on the returns on loans. Therefore, with constant prices, the equity of regulated banks becomes relatively more attractive, inducing households to invest in regulated banks. With binding capital requirements, regulated banks start issuing more debt. In general equilibrium, this leads to higher deposit rates, which increase the relative costs of regulated banks, pushing down their net worth. Lower net worth amplifies excessive risk-taking incentives beyond those which do not consider the interactions with shadow banks. I quantify the relevance of this channel in the numerical part.

Capital requirements affect risk taking through a change in ε^*_{t+1}. When γ_t increases, ε^*_{t+1} falls. The value of limited liability decreases, forcing banks to keep more "skin in the game".

5 Calibration, Estimation, and Numerical Methods

I map the structural parameters to quarterly US data from 1990 Q1 to 2022 Q1 using a combination of calibration and estimation. Table 2 shows the parameter values.

I divide the calibrated parameters into two groups. The first group comprises the parameters that are unrelated to the macro data that the model aims to explain. I assign standard values to them. The parameters in the second group are set to match the first moments of the data with the first moments of selected variables in the steady state of the model.

I use a SMM (simulated method of moments) procedure to pin down the values of the parameters that are sensitive to the second moments. I simulate the economy for 2,000 periods and match empirical moments from HP-filtered data with analogous simulated moments from the model (also HP-filtered). I impose the Ramsey policy for capital requirements in the model.[20] I consider the optimal weighting matrix that minimizes the asymptotic variance of

[20]This assumption is practical for speeding up the SMM procedure. An alternative is to run it through the model with occasionally binding constraints and a constant capital requirement. However, under that approach the algorithm would have to simulate the non-linear model with excessive risk shifting for each proposed combination of values of the estimated parameters, considerably decreasing the computation speed.

the estimates.

5.1 Calibrated Parameters

I choose conventional values for the discount factor β, the capital share α, the intertemporal elasticity of substitution for consumption σ_c, and the depreciation rate δ. The remaining parameters are specific to my framework.

Table 2: Parameters

	Value	Description	
Conventional			
β	0.99	Discount rate	
α	0.3	Capital share in production	
σ_c	1.1	Elasticity of substitution for consumption	
δ	0.025	Depreciation rate	
Calibrated (first-order moments matching with steady-state conditions)			*Target*
τ_R	3.4052	Standard deviation of R-bank idiosyncratic shock (%)	$\frac{\text{Debt}}{\text{EBITDA}} = 7$
ξ_R	0.01	Minus mean of idiosyncratic shock for R-banks (%)	Cap. requirement= 8%
σ_0	0.39	Relative weight on liquidity in the utility function	Real prime rate 2.43% (ann.)
f_R	0.657	Linear Costs of Regulated Banking (%)	$R^l - R^{dR} = 3\%$ (ann.)
f_S	0.3081	Linear Costs of Shadow Banking (%)	$R^{dS} - R^{dR} = 0.15\%$ (ann.)
τ_S	1.8853	Standard deviation of S-bank idiosyncratic shock (%)	Corp. bond default= 1.44% (ann.)
α_1	0.5646	Weight on S-bank deposits in the liquidity function	Share of SB assets= 45%
θ	0.9	S-banker's survival probability	10% dividend payout of SB
λ	0.2295	Fraction of capital that can be diverted by S-banks	Shadow bank leverage= 25
$\underline{\sigma}^R$	0.01	Minimum risk that R-banks can take	numerical solution method
$\bar{\sigma}^R$	0.99	Maximum risk that R-banks can take	numerical solution method
$\underline{\sigma}^S$	0.01	Minimum risk that S-banks can take	numerical solution method
$\bar{\sigma}^S$	0.99	Maximum risk that S-banks can take	numerical solution method
Estimated (SMM procedure)			
σ_d	-2.1975	Substitution elasticity b/w R- and S-bank liquidity, $\frac{1}{1-\sigma_d}$	
ζ	1.3374	Interest rate elasticity of supply of total liquidity, $\frac{1}{\zeta}$	
ϕ	6.1126	Investment adjustment costs	
ε_A	0.9887	Standard deviation of TFP shock (%)	
ε_ι	0.8872	Standard deviation of capital quality shock (%)	
ρ_A	0.98	Persistence of TFP shock (%)	
ρ_ι	0.4865	Persistence of capital quality shock (%)	

Note: See Section 5 for the strategy of mapping the model to the data.

The model is constructed to trace the effects of shocks on optimal capital requirements, and it is not suitable for computing the optimal steady-state value. Basel III stipulates that regulated banks maintain at least 7% of Tier 1 equity relative to risk-weighted assets

To test the significance of my assumption, I calculate the model moments in the model with risk shifting using the estimated parameter values computed for the Ramsey policy. I find that the model moments are close to each other across the alternative procedures. This result supports my assumption.

starting from 2019.[21] Regulated banks usually hold a buffer over this threshold. I calculate the sample period average of the ratio of Tier 1 equity to total assets from the table of assets and liabilities of FDIC-insured commercial banks and savings institutions. It results in the value of 8% over the sample period. I set the steady-state capital requirement at the minimum level to support the safe equilibrium and match it with 8%. This procedure finds the pseudo-optimal static capital requirement such that a small decrease in capital requirements makes banks finance socially sub-optimal projects, while a small increase results in liquidity losses without changing the risk-taking profile.

Two parameters, τ_R and ξ_R, which enter the idiosyncratic process of the risky technology, support a wide range of steady-state capital requirements. I associate the standard deviation of the idiosyncratic shock τ_R with leveraged lending.[22] It is related to financing corporations with high leverage – defined as those with a debt-to-EBITDA ratio greater than 6. To motivate this number, I refer to guidance on the threshold of total debt to EBITDA that the U.S. regulators use for high risk lending to raise supervisory concerns.[23] Conditional on τ_R, I choose ξ_R, which is interpreted as the average loss from financing risky projects, to hit a 8% level of the capital requirement in the steady state.

To match the data on interest rate spreads, I introduce costs of banking. These costs include operating expenses which can be associated with the provision of loans. In particular, each period the bank of type $i = R, S$ pays $F_t^i = f^i l_t^i$ from its current profits.[24] The Appendix provides the details of the effects of these banking costs on the expressions of portfolio returns. The data counterparts for the regulated bank deposit rate and the shadow bank money market borrowing rate are the national rate on non-jumbo deposits and the 3-month financial regulated paper interest rate, respectively. I calibrate σ_0 to hit the real bank prime rate which I use as a measure of the return on safe loans.

I set τ_S to match the default rate on corporate bonds that I borrow from Begenau and Landvoigt (2021). I map the utility weight on shadow bank deposits, α_1, to target the share of shadow bank assets. I use the estimates of the size of the U.S. shadow banking sector from the Financial Stability Board and divide it by the total bank assets from the Board of Governors of the Federal Reserve System. I set $\theta = 0.9$ that implies a shadow bank dividend payout of 10% that is considered in Ferrante (2018). I calibrate λ, the fraction of capital that can be diverted by shadow banks, to hit the shadow bank leverage of 25 which is in line with the leveraged ratios of broker-dealers reported in the literature. The ratio of the leverage of

[21]The ratio was 6% before 2019.

[22]Appendix I describes in detail the choice of τ_R.

[23]See Financial Stability Report (2019).

[24]Introduction of the cost of banking improves the model's fit in the steady state. But it does not affect the main mechanism that I explore in the model.

the two types of banks corresponds to the empirical estimates that the leverage of the shadow banking system is about three times as large as the one of depository institutions obtained from the Flow of Funds by Ferrante (2018).[25] Finally, I consider the share of risky projects within a broad range of $[0.01, 0.99]$ for both types of banks. This choice makes the minimum scale assumption technical, and thus it does not contaminate the model's main mechanisms.

5.2 Estimated Parameters

I estimate the remaining parameters, namely, the elasticity of substitution between regulated and shadow bank liabilities, σ_d, the interest rate elasticity of supply of total liquidity, ζ, the investment adjustment cost, ϕ, and the parameters governing AR(1) processes of technology and capital quality shocks that drive the economy in my model. The target moments are second moments (variances, correlations, and auto-correlations) of HP-filtered natural logarithms of real output, investment price, real RB liabilities, and real SB liabilities multiplied by 100.

The curvature parameters σ_d and ζ in the utility function determine how much RB and SB liabilities can vary in the model. For this reason, I target the second moments of the two types of liabilities in the data. The investment adjustment cost parameter governs the dynamics of investment price. The parameters capturing the shock processes restrict the behavior of real GDP and investment price. Their values are governed by the second moments of these variables together with the correlations with other variables selected for the moment-matching exercise.

Table 3 shows that the model moments are close to the data moments. Importantly, the variances of real RB and SB liabilities are matched very well. The signs of correlation coefficients of RB and SB liabilities with output are in line with the stylized facts outlined in Section 2. At the same time, the model overestimates the correlation of SB liabilities with GDP.[26] The model also over-predicts the variance of investment price. The model does a good job at matching auto-correlations.

5.2.1 Discussion of Calibrated Parameters

The value of the investment adjustment cost, ϕ, is important for the volatility of investment price that is overestimated by my model relative to the data. Smaller values of ϕ would make

[25] This choice contrasts with Jiang et al. (2020) who find that shadow banks are less leveraged than regulated banks in the mortgage origination business using call report data. I focus on excessive risk and have in mind opaque practices which are usually associated with non-regulated participants. The market of leveraged loans is a case in point.

[26] I find that if we truncate our sample period to 2019 Q4 (to the period before the Covid shock), then the empirical correlation coefficient goes up to 0.5, making it closer to the corresponding theoretical moment.

Table 3: Matching Moments

	Data	Model
Var(GDP)	1.89	1.95
Corr(GDP, Investment Price)	0.11	0.74
Corr(GDP, RB Liabilities)	-0.14	-0.09
Corr(GDP, SB Liabilities)	0.26	0.84
Var(Investment Price)	0.66	1.62
Corr(Investment Price, RB Liabilities)	0.31	-0.36
Corr(Investment Price, SB Liabilities)	0.40	0.79
Var(RB Liabilities)	3.16	3.28
Corr(RB Liabilities, SB Liabilities)	0.41	-0.07
Var(SB Liabilities)	19.19	19.19
Autocorr(GDP)	0.63	0.74
Autocorr(Investment Price)	0.72	0.57
Autocorr(RB Liabilities)	0.76	0.59
Autocorr(SB Liabilities)	0.91	0.69
Loss	2.0181	

Note: This table reports the variance, correlation, and auto-correlation of variables in the model and in the data. All these variables are used in the SMM estimation. The data and model-simulated time series are logged, then HP filtered, and multiplied by 100. Let M denote moments and W denote the optimal weighting matrix that minimizes the variance of the estimates. Then Loss $= (M_{model}(\theta) - M_{data})' \, W \, (M_{model}(\theta) - M_{data})$, where θ is a vector of estimated parameters.

the model moments associated with investment price closer to their empirical counterparts keeping the values of all other estimated parameters unchanged.[27] However, smaller values of ϕ would significantly affect the volatility of RB and SB liabilities as well as their correlations with real GDP, making them inconsistent with the stylized facts described in Section 2. Since I construct my model to explain the behavior of regulated and shadow banks, I find it more expedient to focus on the performance of the model moments associated with the banking side of the economy. The estimated value of ϕ delivers reasonable results on this domain.

Table 4 reports the performance of the model regarding some other key variables that I do not target in the SMM estimation. The idea is to evaluate how my model can speak about other important features of the data that are not explicitly taken into account when estimating the parameters of the model.

Without being moment-matching targets, the model captures the correlation of investment with other considered variables well. Investment is chosen on the grounds of being related to investment price. While the model overestimates the negative correlation between the size

[27]These model moments include volatility of investment price and cross-correlations of investment price with other variables selected in the moment-matching exercise.

of RB relative to SB liabilities and real GDP and underestimates the positive correlation between the size of RB relative to SB liabilities and the credit spread, it captures the signs of both correlations correctly. These results indicate that the model's mechanisms implied by the calibrated values do not contradict the second and third stylized facts presented in Section 2.

Table 4: Untargeted Correlations

	Data	Model
Corr(Investment, GDP)	0.83	0.69
Corr(Investment, Investment Price)	0.13	0.23
Corr(Investment, RB Liabilities)	0.05	0.01
Corr(Investment, SB Liabilities)	0.26	0.19
Autocorr(Investment)	0.82	0.93
Corr(Difference b/w RB and SB Liabilities, GDP)	-0.35	-0.79
Corr(Difference b/w RB and SB Liabilities, Credit Spread)*	0.3	0.09

Note: This table reports the correlation and auto-correlation of variables in the model and in the data. The data and model-simulated time series are logged, then HP filtered and multiplied by 100. *Real RB and SB liabilities are measured by year-ahead deviations from their corresponding HP trends to be consistent with the procedure used in Section 2. The credit spread in the model is calculated by $E_t(R^l_{t+1}) - R^{dR}$.

5.3 Numerical Methods

In Section 6.1, I solve the model using local perturbation methods. Sections 6.2 and 7 require nonlinear methods to account for the endogenous transition between the regimes of different risk-taking depending on the state vector. I solve the model by applying the OccBin toolkit developed in Guerrieri and Iacoviello (2015). OccBin modifies a first-order perturbation method and employs a guess-and-verify approach to obtain a piecewise linear solution under perfect foresight.[28] The algorithm has advantages over nonlinear projection methods because it is computationally fast and applies to nonlinear models with a large number of state variables. Appendix J.2 describes the adaptation of my framework to Occbin.

6 Numerical Results

In this section, I conduct a quantitative analysis to estimate the impact of shadow banks on the design of optimal policies. Section 6.1 shows that the interactions of regulated and shadow

[28] The reader should also be aware that OccBin solution does not capture precautionary behavior linked to the possibility of moving away from the reference regime in the future.

banks can change the magnitude and direction of optimal responses of capital requirements to standard business-cycle shocks. Section 6.2 evaluates how tighter capital regulation affects the economy that starts at suboptimally low capital requirements. A better designed policy that calls for relatively higher capital requirements can lead to the migration of credit from shadow banks to regulated banks.

6.1 Impact of Shadow Banks on Optimal Dynamic Policies

I quantify the relevance of the interaction of regulated and shadow banks by comparing the optimal response of capital requirements to shocks in two models: the baseline model with shadow banks and the model with regulated banks only. I propose a framework with regulated banks only which follows the presentation of the model economy but excludes shadow banks. In this regard, the regulated banking sector intermediates all the assets, and the weight on regulated bank liquidity α_1 is one. Thus, the liquidity preferences transform into

$$\Psi_2\left(D_t^R\right) = \frac{\left(\left(D_t^R\right)^{\sigma_d}\right)^{\frac{1-\zeta}{\sigma_d}}}{1-\zeta} = \frac{\left(D_t^R\right)^{1-\zeta}}{1-\zeta} \tag{30}$$

in the model with regulated banks only. All other structural parameters are unchanged. The steady-state capital requirements and the interest rates are not affected. The only difference is the absence of the reintermediation channel.

To compute optimal capital requirements, I consider the Ramsey problem in which the planner chooses the path of capital requirements to maximize the conditional expectation of the household's utility as of time zero subject to the restrictions of the decentralized equilibrium. Appendix J.1 shows that the solution to this program is to set the optimal capital requirement at the lowest level that prevents excessive risk taking in every period following each shock. In the language of the model, it is the minimum capital requirement that supports the safe equilibrium. Intuitively, any capital requirement above this level would result in liquidity losses, while any level below this level would be suboptimal due to financing socially inefficient risky projects.

Figure 4 summarizes the results on the response of optimal capital requirements to business-cycle shocks. It plots the peak in the increase or trough in the decrease of optimal capital requirements for a given initial change in either technology or capital quality. It compares the responses to these shocks for the full model with shadow banks (the solid line) to a restricted model with only regulated banks (the dashed line). The values on the horizontal axis denote the corresponding change in output. For TFP shocks, the two lines slope downwards but the slope of the solid line is larger, meaning that the optimal capital requirement reacts more aggressively once shadow banks are taken into account. For capital quality shocks, the two lines have different slopes, meaning that capital quality shocks call for a different direction of a change in the optimal capital requirement once shadow banks are considered. The next two subsections illustrate the mechanisms of the optimal response of capital requirements to each of these two shocks.

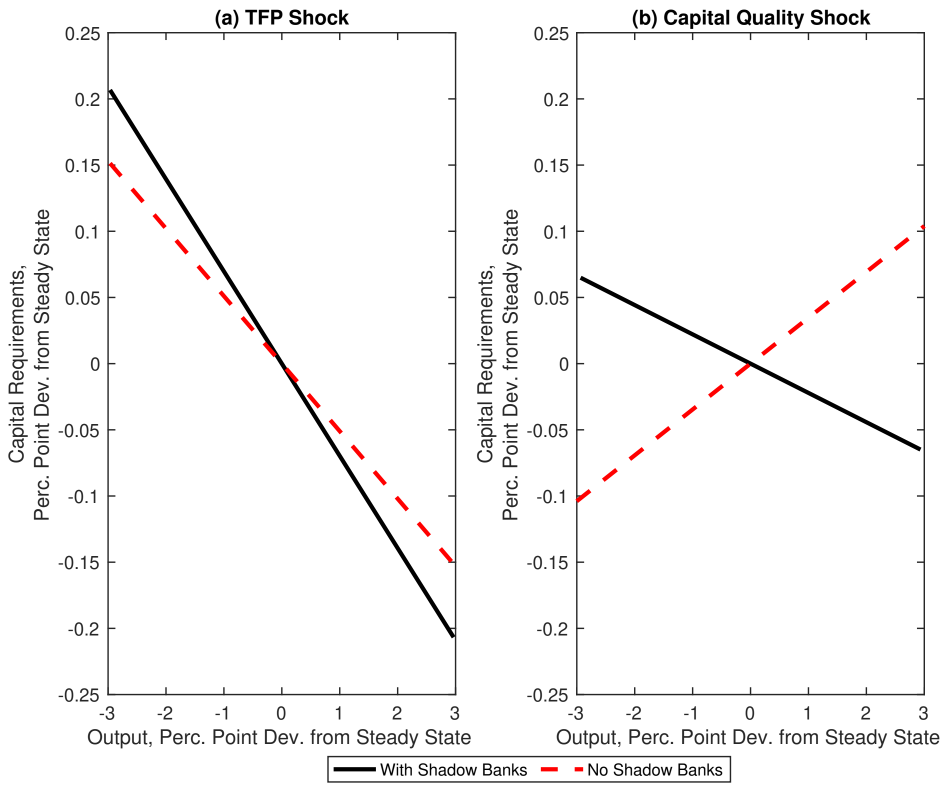

Figure 4. Accounting for Shadow Banks Affects (a) the Magnitude and (b) the Direction of the Optimal Reaction of Capital Requirements.

Note: This figure plots the maximum responses (in absolute value) of optimal capital requirements to (a) a TFP shock and (b) an unexpected improvement in the productivity of installed capital – a capital quality shock à la Gertler and Karadi (2011) – for two models. The solid line shows the responses of the baseline model that takes into account the interactions between regulated banks and shadow banks. The dashed line represents what would happen in the model without such interactions (regulated banks only). The vector of shock sizes is the same for each of the two models. The values on the abscissae denote the corresponding change in output.

6.1.1 TFP Shock: Magnification of the Response of Capital Requirements

Figure 5 shows the effects of a contractionary TFP shock. The shock decreases A_t by one standard deviation and follows an AR(1) process with the estimated persistence parameter. The solid lines represent the responses of the economy to the shock under the optimal capital requirements set in the baseline model that includes both regulated and shadow banks. The

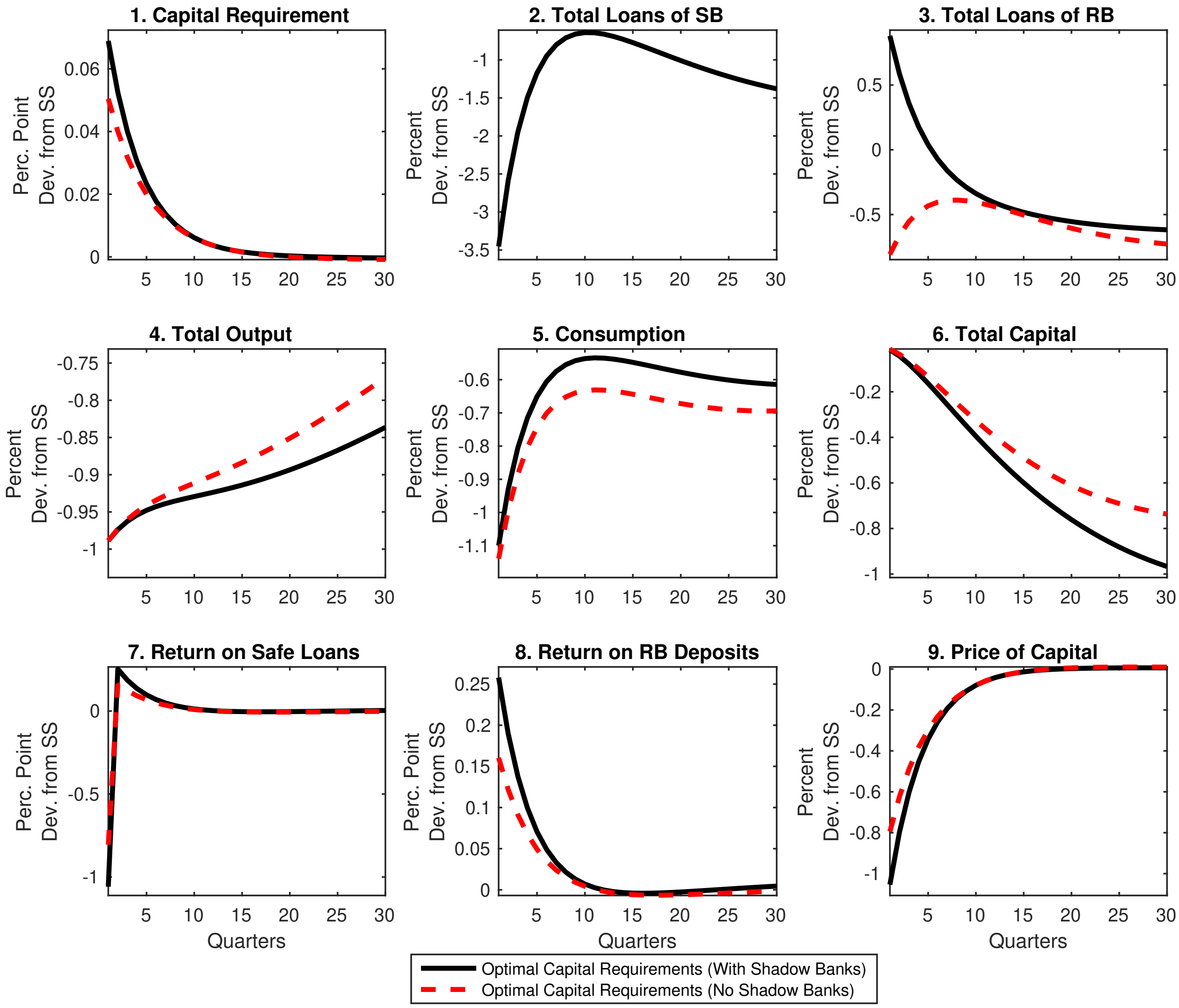

Figure 5. A Contractionary TFP Shock: The Optimal Capital Requirement Responds More Aggressively Once Shadow Banks are Taken Into Account.

Note: This figure plots the responses of the model's variables (in deviation from the steady state) to a one s.d. fall in A_t under optimal capital requirement (set at the minimum level to prevent excessive risk taking in every period following the shock) for two models. The shock follows the estimated AR(1) process. The solid line shows the responses of the baseline model with shadow banks. The dashed line represents what would happen in the model without shadow banks (regulated banks only). SB and RB stand for shadow and regulated banks, respectively. SS means steady state.

dashed lines show the responses of the economy to the same shock under the optimal capital requirements set in the model that includes regulated banks only.

Panel 1 in Figure 5 illustrates that optimal capital requirements react more strongly to the TFP shock once shadow banks are taken into account. The same panel shows that the optimal capital requirement is, on impact, around 0.02 percentage point greater than it is under the optimal policy that disregards shadow banks. This difference amounts to about 30 percent of the total increase in optimal capital requirements.

What are the mechanisms? The TFP shock reduces the returns on loan portfolios of banks. This fall decreases the net worth and launches the familiar financial accelerator mechanism described by conventional models of financial frictions: a drop in the net worth of banks increases agency costs, forcing banks to sell their assets, thus depressing asset prices and further worsening their balance sheet conditions.

There is one element that stands out. Unless there are shadow banks, regulated banks decrease their demand for loans as capital becomes less productive, so their loans fall (the dashed line in Panel 3). By contrast, the loans migrate from shadow banks (the solid line in Panel 3) into regulated banks (the solid line in Panel 2) in the model with shadow banks. In fact, the returns on regulated bank equity are relatively less negatively affected than the corresponding returns of more leveraged shadow banks. This feature makes the equity of regulated banks relatively more attractive for households. Consequently, with binding capital requirements, regulated banks start demanding more loans and deposits. Since the two types of deposits are imperfect substitutes, households require a higher deposit rate (Panel 8) to substitute regulated bank deposits for shadow bank deposits. This reintermediation further decreases the loan returns of regulated banks (Panel 7), making risk more attractive. Although the higher capital requirement pushes down profitability of regulated bank equity, it eliminates socially inefficient risk and thus makes the business-cycle variables, such as output (Panel 4), consumption (Panel 5), and total capital (Panel 6) less affected by the negative shock.

The takeaway is that shadow banks magnify the impact of shocks, such as TFP shocks that have been found to be important drivers of business cycles, on optimal capital requirements.[29]

6.1.2 Capital Quality Shock: A Change in the Direction of Capital Requirements

Here I consider an alternative business cycle shock which is commonly used in the literature on financial frictions. Let ι_t denote the quality of capital. At the beginning of each period, one unit of capital transforms into ι_t units of effective capital used in production. Appendix

[29] See Galí (1999) for the empirical evidence on the effects of technology shocks.

K specifies the details on the inclusion of the capital quality shock into the problems of banks and firms. The quality of capital now provides additional variation in the returns on safe and risky loans, affecting the risk-taking incentives of banks.

I hit the economy with a positive capital quality shock that follows the estimated AR(1) process. I fix the size of the shock so that it leads to the same percentage change (in absolute value) of output, on impact, as for the TFP shock. The results are reported in Figure 6. As before, in each panel, the solid line represents the responses of the baseline model, while the dashed line shows what the optimal capital requirements would be in the absence of shadow banks.

Panel 1 shows that the optimal capital requirement falls in the model with shadow banks, while it rises in the model with regulated banks only. The positive capital quality shock pushes up the returns on safe loans. It makes households richer, so the wealth effect expands consumption. The substitution effect makes investment more attractive by taking advantage of the positive shock before it dissipates. The balance between these two effects determines the direction of the response of capital requirements.

Panel 7 illustrates that the safe returns increase by more in the model with shadow banks. In fact, the higher leverage ratio of shadow banks magnifies the effects of the positive shock on their equity returns, making shadow bank equity more attractive in expectation.[30] This force allows shadow banks to attract additional funds from households and finance more investment by issuing more loans (Panel 2) which are also partially coming from loans of regulated banks (Panel 3). The substitution effect gets a boost. In the presence of adjustment costs, the higher demand for investment pushes up the price of installed capital (Panel 9), increasing capital gains. This force becomes so strong as it increases the returns on safe projects for regulated banks, and thus safe projects become more attractive. Capital requirements fall.

This reallocation of equity increases the intermediation of loans by shadow banks (Panel 2) also coming from regulated banks (the solid line in Panel 3). Shadow banks are able to attract more funds from households, increasing the relevance of the substitution effect. Thus, the demand for investment increases in the model with shadow banks, so the price of capital rises more (Panel 9), making capital gains larger. Higher loan returns lead to a fall in risk-taking incentives and thus provide some leeway for the economy to benefit from greater liquidity if the capital requirement falls depicted by the solid line in Panel 1. As the shock dissipates and the capital stock remains elevated, the marginal product of capital decreases, and thus safe projects become less attractive.

[30] The described dividend payout policy of shadow banks, which is inherited from their constant exit rate, weakens the initial effect of the positive shock on shadow bank equity. Moreover, it will be shown that the capital requirement falls, so the equity of regulated banks becomes attractive, ensuring some initial inflow into regulated bank equity, on impact.

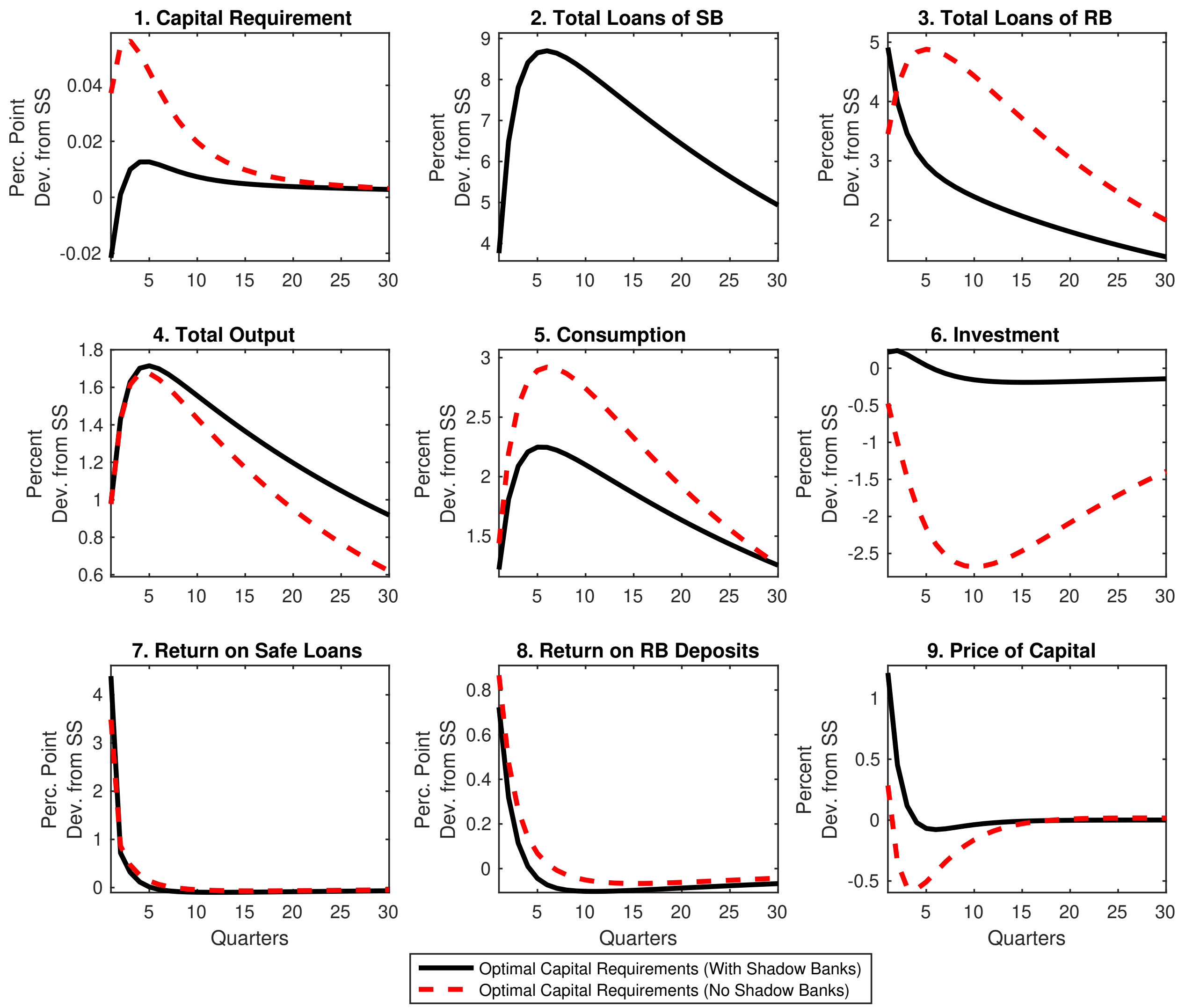

Figure 6. An Expansionary Capital Quality Shock: The Optimal Capital Requirement Changes its Direction Once Shadow Banks are Taken Into Account.

Note: This figure plots the responses of the model's variables (in deviation from the steady state) to a rise in the capital quality, ι_t, under optimal capital requirement (set at the minimum level to prevent excessive risk taking in every period following the shock) for two models. The shock follows the estimated AR(1) process. The shock is sized to lead to the same percentage change of output, on impact, as in the case of the TFP shock. The solid line shows the responses of the baseline model with shadow banks. The dashed line represents what would happen in the model without shadow banks (regulated banks only). SB and RB stand for shadow and regulated banks, respectively. SS means steady state.

By contrast, the expansionary capital quality shock increases the risk-taking incentives of regulated banks in the model without shadow banks. In fact, the investment adjustment costs curb the substitution effect, making the wealth effect relatively stronger. Investment falls (Panel 6) and total output (Panel 4) increases by less in the model that disregards shadow banks. Consumption expands by more (Panel 5) because of smaller investment. There is a boost to the profits that comes from the increase in the capital quality but since investment goes down, the price of installed capital is decreasing. Thus, smaller capital gains make the safe returns relatively lower, supporting the result that the substitution effect is weaker. The shock increases the capital stock. As the shock dissipates and the capital stock remains elevated, the marginal product of capital falls, and thus safe projects become less attractive, justifying higher capital requirements in the environment where, in fact, lower capital requirements are needed.

6.2 Effects of Higher Capital Requirements: Tighter Regulation Can Lead to Greater Intermediation of Credit by RB

To evaluate the post-crisis measures on strengthening the balance sheets of regulated banks, I compare the impact of tighter regulation on the economy depending on the magnitude of a rise in capital requirements. I use exactly the same values of all structural parameters as in Section 5. I consider an increase in capital requirements starting from a suboptimally low level of 6% in the non-stochastic steady state. This level leads to excessive risk taking that is associated with inefficient lending.[31] Then I measure the effects of two policies that persistently increase the capital requirement but differ in the magnitude of its change.

Figure 7 plots the results of this experiment. The solid lines show the dynamic responses of the economy under the policy that achieves the optimal level of the capital requirement that is just enough to overturn the financing of socially inefficient projects.[32] The dashed lines represent the dynamic responses of the economy in which there is an insufficient increase (by 1%) in the capital requirement. There are several observations.

First, the benefits from higher capital requirements depend on the magnitude of the change. Panel 2 in Figure 7 shows that the insufficient increase in the capital requirement does not move the economy away from financing socially inefficient risky projects compared to its optimal increase. This fact leads to the differences in the responses of consumption,

[31]Increasing capital requirements from the optimal level would have a different result but it would be mixing up the fact that this would be a suboptimal increase. I am interested in characterizing an optimal increase.

[32]Pushing capital requirements above this optimal level does not affect risk taking but it reduces households' welfare by making banks less reliant on deposits, which are valued by households for their liquidity.

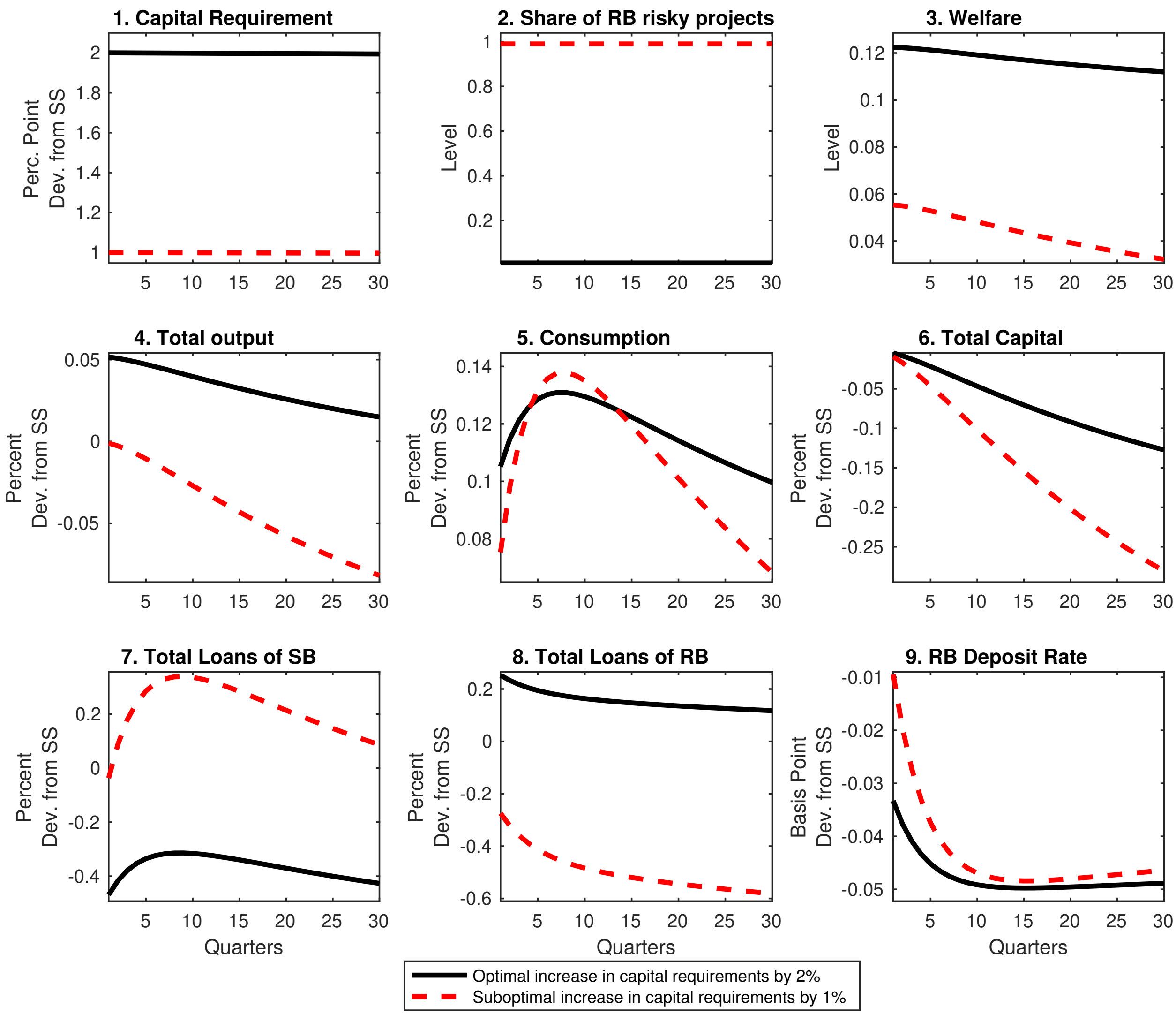

Figure 7. An Increase in the Capital Requirement Starting From a Suboptimally Low Level: The Better Policy Leads to Greater Reintermediation of Credit by Regulated Banks.

Note: This figure plots the responses of the model's variables (in deviation from the steady state) to an increase in the capital requirement starting from the suboptimally low steady-state level of the capital requirement at 6% (risky equilibrium). The solid line shows the dynamic responses of the economy that achieves the optimal level of the capital requirement. The dashed line represents the dynamic responses of the economy in which there is an insufficient increase in the capital requirement. The shocks follow the same AR(1) process with an autoregressive coefficient of 0.9999. SB and RB stand for shadow and regulated banks, respectively. SS means steady state.

output, total capital and loans.

Second, higher capital requirements make regulated bank deposits scarce. Liquidity becomes relatively more expensive, so households prefer to substitute away from liquidity services to consumption. Therefore, in contrast to the mechanism described so far that pushes up the deposit rates, here migration of credit does not result in higher borrowing costs of regulated banks. Panel 9 shows that the deposit rate falls. Moreover, since intermediation becomes more costly, banks lend less, so both capital and investment drop. The fall in capital stock raises the marginal product of capital. This force also makes the safe projects more attractive.

Third, although tighter capital requirements make regulated banks less profitable, this fact does not necessarily cause contraction of credit of regulated banks and expansion of loans of shadow banks. Panels 7 and 8 show that the optimal increase in the capital requirement leads to migration of credit from shadow banks to regulated banks, while the suboptimal increase in the capital requirement has the opposite effect on it. In fact, the deposit rate of regulated banks falls more for the optimal rise in the capital requirement, which is higher, because deposits are relatively more scarce (the convenience yield is higher). A lower deposit rate decreases the costs of providing loans by regulated banks. This mechanism increases lending of regulated banks.[33] In contrast, the suboptimal increase in the capital requirements is not enough for fully activating this mechanism, so it makes regulated banks relatively less profitable and thus causes expansion in shadow banking.

Finally, the suboptimal increase in capital requirements can do more harm rather than good. In fact, total capital is decreasing (the dashed line in Panel 6), and output is falling (the dashed line in Panel 4). Panel 5 illustrates that consumption is hump-shaped: initially, it is increasing but then starts falling. In contrast, the optimal rise in capital requirements enables the economy to benefit from the more efficient technology. Consumption increases by more (the solid line in Panel 5) and output goes up (the solid line in Panel 4), leading to higher welfare (Panel 3).

The takeaway is that higher capital requirements can lead to the migration of credit from shadow banks to regulated banks as a result of a better designed capital regulation policy.

[33]Begenau (2020) also finds that higher capital requirements can decrease the cost of banking. My model extends this result to the environment with both regulated and unregulated banks where the migration of credit is possible across the two types of banks.

7 Implementing the Optimal Dynamic Capital Requirements

It seems unrealistic to expect that bank regulators can distinguish a TFP shock from a capital quality shock in real time, and therefore formulate and enforce a change in capital requirements tailored to the type of shock. In practice, policymakers can devise implementable rules. I study the ability of simple rules to mimic the response of the Ramsey optimal policy.

I choose a rule in which capital requirements respond to changes in the loan-to-output ratio, in line with the Basel III regulation. I use data generated by the simulations for 500 periods. I regress the Ramsey capital requirement on a constant and the logarithm of the loan-to-output ratio. I consider two specifications.

$$\text{Cap\_Req}_t = b_0 + b_R \log\left(\frac{\text{RB Loans}_t}{\text{GDP}_t}\right) + b_S \log\left(\frac{\text{SB Loans}_t}{\text{GDP}_t}\right), \tag{31}$$

$$\text{Cap\_Req}_t = b_0' + b_{tot}' \log\left(\frac{\text{Total Loans}_t}{\text{GDP}_t}\right). \tag{32}$$

In the first specification (eq. (31)), I differentiate between two types of loans provided by each of the two banking sectors. I call it the proposed rule. In the second specification (eq. (32)) that relates to the Basel III guidance, I consider total loans ($\text{Total Loans}_t = \text{RB Loans}_t + \text{SB Loans}_t$). I dub it the Basel rule. I then compare the results against various performance measures. One measure is an R-squared of the regression. It indicates how closely a rule tracks the Ramsey setting.

Table 5 reports the results. The proposed rule delivers the R-squared of 0.725. The capital requirement responds positively to the ratio of RB loans to GDP but negatively to the ratio of SB loans to GDP. Intuitively, policymakers impose tighter capital regulation when a) excessive risk taking is accumulating due to the fact that RBs intermediate more credit and b) loans are migrating from SBs to RBs. The latter activates the risk-taking mechanism coming from the interactions of the two types of banks. The reaction of the optimal capital requirement to the ratio of RB loans to GDP is about one order of magnitude larger than its response to the ratio of SB loans to GDP. The implication is that the capital requirement changes more aggressively to the developments in regulated banks, i.e. the sector for which the instrument is specifically designed.

To interpret the regression numbers, I calculate the standard deviations of the simulated data on the loan-to-output ratios and plug them into the regression for the proposed rule. In particular, a one standard deviation increase in the RB credit-to-GDP gap calls for a 0.1% rise in capital requirements, while a one standard deviation increase in the SB credit-to-GDP

Table 5: **Basel-III Style Simple Rules**

	Proposed Rule (eq. (31))	Basel Rule (eq. (32))
(Intercept)	0.0611*** (0.0006)	0.0749*** (0.0007)
$\log\left(\frac{\text{RB Loans}_t}{\text{GDP}_t}\right)$	0.004*** (0.0001)	
$\log\left(\frac{\text{SB Loans}_t}{\text{GDP}_t}\right)$	-0.0007*** (0.0001)	
$\log\left(\frac{\text{Total Loans}_t}{\text{GDP}_t}\right)$		0.0005*** (0.0001)
R-squared:	0.725	0.0724
Loss of annual consumption (Δ):	0.091%	0.1585%
Share of periods with excessive risk taking:	39.4%	57.9%

Note:*p<0.10,**p<0.05,***p<0.01 This table reports the results on the regression of the Ramsey optimal capital requirement on the logarithm of the loan-to-output ratio for two specifications. The data is generated by simulating the model with the Ramsey optimal capital requirement in place for 500 periods. The specification with the proposed rule differentiates between RB and SB loans. The specification with the Basel rule aggregates loans across the two banking sectors. The welfare loss, Δ, is consumption equivalent cost of the suboptimal economies relative to the Ramsey economy under the unconditional welfare metric.

gap leads to a 0.03% fall in capital requirements.

In contrast, the Basel rule does not work well; its R-squared is 0.0724. The capital requirement responds positively to the ratio of total loans to GDP. But this ratio is not informative to guide policymakers about capital requirements. It does not use information about the composition of credit that determines excessive risk taking through the interactions of banks. Several related papers that place no role for shadow banks in their frameworks also find a low R-squared of the Basel rule in a similar regression analysis.[34]

7.1 Welfare Measure

I solve the model under each of the proposed rules using Occbin. I simulate the economy 4 times for 500 periods, and I calculate average values of different statistics on the performance of the rules over the simulations.

I measure welfare loss under the proposed and Basel rules relative to the Ramsey optimal policy. I consider unconditional welfare as a metric. I compute consumption-equivalent

[34]For example, the R-squared of the Basel rule equals 0.016 in Canzoneri et al. (2020). Davydiuk (2018) finds a value of 0.08.

welfare losses by sizing the permanent tax in annual consumption in the Ramsey optimal economy, Δ, required to make the household as well off as in the economy with the proposed policy rules. Specifically,

$$\begin{aligned} E\sum_{t=0}^{\infty}\beta^t\left[\frac{((1-\Delta)C_{\text{opt},t})^{1-\sigma_c}-1}{1-\sigma_c}+\sigma_0\Psi\left(D^R_{\text{opt},t},D^S_{\text{opt},t}\right)\right]= \\ E\sum_{t=0}^{\infty}\beta^t\left[\frac{C^{1-\sigma_c}_{\text{rule},t}-1}{1-\sigma_c}+\sigma_0\Psi\left(D^R_{\text{rule}_j,t},D^S_{\text{rule}_j,t}\right)\right], \end{aligned} \tag{33}$$

where the subscript "opt" denotes the Ramsey optimal policy and the subscript "rule" indicates one of the rules depending on $j=\{\text{Proposed, Basel}\}$. The Appendix shows the derivation of Δ.

Table 5 shows the computed values of Δ for each of the rules. The proposed rule is about twice less costly in terms of the permanent consumption than the Basel rule. This result further supports the better performance of the rule that differentiates loans across the banking sectors.

7.2 Capital Buffers

I consider the share of periods with excessive risk taking to explore the power of the rules to offset inefficient lending. Although the proposed rule mimics the Ramsey optimal capital requirements rather well in the R-squared metric, it leads to excessive risk taking in about 40% of time. The Basel rule performs worse, amounting to around 60% of periods of inefficient lending.

A relatively high share of excessive risk-taking episodes resulted from the rule that mimics the Ramsey policy well in the R-squared metric can be explained as follows. While the R-squared is high, the proposed rule does not perfectly match the Ramsey policy. So it still makes possible a situation in which the rule does not prevent excessive risk taking. The Ramsey policy is constructed to always offset inefficient lending, so excessive risk taking occurs off its equilibrium path. Thus, the rule that mimics the Ramsey policy under the conditions that excessive risk taking is impossible does not necessarily work well once the economy moves into excessive risk taking. For example, it may not react sufficiently enough to return the economy back to the safe equilibrium. In fact, the inspection of the simulated series confirms this possibility. There is usually a sequence of successive risk-taking episodes. Some of these sequences can last for more than 30 consecutive periods. A larger increase in the capital requirement could immediately prevent excessive risk taking but this information is not available to the rule. Another possibility is to require banks to hold extra capital in the form of buffers.

This analysis motivates the study of combining simple rules and static capital buffers. A buffer is extra capital that increases static capital requirements. I consider different values for the buffer, while keeping all structural parameters unchanged. In this regard, capital buffers provide some cushion against potential losses from shocks that could move the economy into excessive risk taking. However, they are costly in the steady state because they are imposed above the optimal level of the steady-state capital requirement.

Figure 8 shows how jointly combining the capital buffer with each of the rules affects the consumption equivalent loss, Δ. It plots the loss over different parameter values of the capital buffer. The red dashed and green solid lines represent the Basel and proposed rule from the specifications in Table 5, respectively. I also compare their performance against imposing the buffer only (the blue dotted line).

First, the blue dotted line lies above the red dashed and green solid lines. So combining capital buffers and the rules welfare dominates imposing the buffers only. The policy implication is that devising simple rules can be useful.

Second, the lines of Figure 8 have a convex shape. The losses are falling for relatively small values of the buffer but then they start increasing. Intuitively, smaller values of the buffer are not enough to prevent excessive risk taking, while larger buffers are inefficient because they decrease liquidity services without affecting risk taking. So it is possible to find the global minimum for each of the lines.

Third, the proposed rule performs relatively better for smaller buffers. Adding around 18 basis points to its specification minimizes the consumption equivalent loss. Intuitively, this small buffer solves the mentioned possibility of being stuck in the excessive risk-taking regime for many consecutive periods, and it still applies knowledge about the optimal dynamic response due to its relatively high R-squared. However, the Basel rule that includes a capital buffer of about 27 basis points works even better. It achieves the smallest loss.

The result of a better performance of the rule with a low R-squared may sound counter-intuitive. It can be explained as follows. The Basel rule with a buffer works worse than the proposed rule for smaller values of buffers because they are not enough to counteract excessive risk taking. Larger buffers can completely shut down the inefficient lending channel that comes from the interaction of banks. It makes the differentiation between the types of loans relatively less useful and sometimes harmful. For example, when the loan-to-GDP ratio falls, the buffer can become excessive relative to the risk-taking incentives that the economy is currently facing. In fact, the efficiency calls for a decrease in the buffer. However, the proposed rule puts significantly more weight on the developments in the RB sector, so it misses the possibility that a fall in the loan-to-output ratio comes from a decrease in SB loans rather than a decrease in RB loans. In fact, the proposed rule can even increase the buffer if there is the migration of credit from shadow banks towards regulated banks. But it is suboptimal given the already relatively excessive value of the capital buffer. The more general Basel rule equipped with capital buffers does not fall into trap of this possibility. It decreases the buffer, making it closer to the Ramsey optimal level.

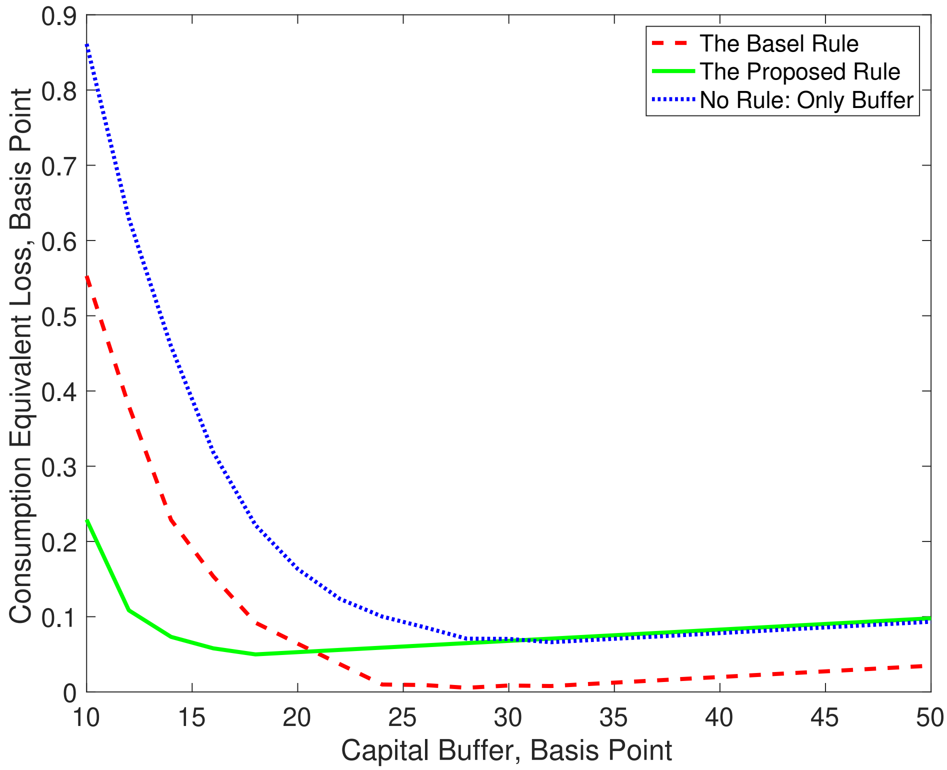

Figure 8. Welfare Losses From Capital Buffers With/Without Basel-III Style Rules

Note: This figure plots the consumption equivalent loss, Δ, depending on the size of the imposed capital buffer. The red dashed line depicts the performance of the Basel rule with the imposed capital buffer on the horizontal axis. The green solid line depicts the performance of the proposed rule with the imposed capital buffer on the horizontal axis. The blue dotted line depicts the performance of the static buffer only. The rules follow specifications in Table 5. The welfare loss, Δ, is consumption equivalent cost of the suboptimal economies relative to the Ramsey economy under the unconditional welfare metric.

8 Discussion

This section discusses two additional results which are not in the primary focus of the paper but may be still useful for understanding the role of the interactions of the two types of banks for policy in my framework. These results, derived from the effects of sectoral shocks, provide further support for the significance of the reintermediation channel for optimal capital requirements.

First, capital requirements can counter the disturbances arising purely in the shadow banking sector, for which policymakers lament the absence of relevant policy tools. For example, these disturbances include a rise in subprime defaults (which set the global financial crisis in motion) or possible losses from holding the portfolios of firms with high debt-to-EBITDA ratio in the market of syndicated loans. The Appendix considers a positive shock to the idiosyncratic standard deviation of risky returns of shadow banks τ_t^S. This sectoral shock increases the quarterly default rate of shadow bank loans. I show that optimal capital requirements rise by about 15 basis points, on impact. The reason for such an increase is the same reintermediation of credit from shadow banks to regulated banks, making regulated banks be affected by the shock. So capital requirements on regulated banks can be used to react to purely sectoral shocks that occur in the shadow banking sector.

Second, not all shocks lead to the reintermediation of credit and thus have implications for capital requirements arising from the interactions of banks. The Appendix considers a positive shock to the idiosyncratic standard deviation of risky returns of regulated banks τ_t^R. The shock increases the deposit-insurance subsidy and calls for a large rise in the optimal capital requirement by nearly 1.5 percentage points, on impact. However, it does not activate the reintermediation channel. The shock has no first-order effect on shadow banks. Higher capital requirements attain both objectives: they eliminate risk-taking incentives and fully stabilize the impact of the shock on aggregate variables.

9 Conclusion

Both academics and policymakers have been calling for greater understanding of the spillover of risk across regulated financial intermediaries and shadow banks, and the implications of such spillovers for capital regulation policies. I consider how the interaction of regulated and shadow banks can affect the optimal path of capital requirements in a quantitative environment with endogenous risk taking. To this end, I estimate a macroeconomic model with the financial sector that includes both regulated and shadow banks. The shield of limited liability and deposit insurance can make socially inefficient projects attractive to banks. Higher capital requirements decrease the benefits from limited liability, counteracting excessive risk taking at the cost of lower liquidity provision.

Three main results stress the importance of tracking both shadow bank and regulated bank lending for the design of optimal capital requirements. First, competition of the two types of banks for funding their loans triggers an endogenous migration of loans and debt, following shocks that depress the safe returns. I find that this channel calls for an additional increase in optimal capital requirements. I provide examples of shocks for which the difference in the optimal response of capital requirements is both in magnitude and in sign.

Second, I discuss the rising concerns of policymakers about migration of risk and loans from regulated banks to shadow banks. I find that optimal tightening of capital requirements can lead to migration of credit from shadow banks to regulated banks because it can decrease costs of providing loans by regulated banks through general equilibrium effects. An insufficient increase in capital requirements can do more harm than good, such as decrease output.

Third, I study how the optimal capital requirement can be implemented in practice. I consider two specifications of Basel-III style rules that relate capital requirements to the loan-to-output ratio. A rule that differentiates loans across the banking sectors outperforms a standard Basel rule that conditions on total loans. While the simple rule that differentiates between loan types can mimic the optimal Ramsey rule much better, it still fails to avoid excessive risk-taking episodes. I show that slightly elevating the steady-state capital requirements in conjunction with the use of these simple rules comes closer to mimicking the performance of the Ramsey optimal rule.

However, both rules fall into risk-taking trap because they do not use the same amount of information available to a Ramsey planner. They generally fail to sufficiently increase capital requirements once the economy enters risk-taking episodes. I show that implementing the considered rules jointly with slightly elevated static capital requirements leads to smaller welfare losses compared to using static buffers only.

These results have clear implications for how regulatory framework could be improved to account for the interactions of regulated and shadow banks. The paper focuses on the role of unregulated shadow banks for the transmission mechanism of capital requirements, and so it abstracts from regulation of shadow banks. An interesting avenue of future research would be to evaluate proposals regarding the regulation of shadow banks featured in the so-called Minneapolis plan.[35] Moreover, exploring the influence of shadow banks on the jointly optimal conduct of monetary and macroprudential policy would be a fruitful area of future research.

[35] See Minneapolis Plan (2017).

References

Akinci, O., & Queralto, A. (2022). Credit spreads, financial crises, and macroprudential policy. *American Economic Journal: Macroeconomics*, *14*, 469–507. https://doi.org/10.1257/MAC.20180059

Basel Committee on Banking Supervision. (2010). Basel committee on banking supervision guidance for national authorities operating the countercyclical capital buffer. https://www.bis.org/publ/bcbs187.pdf

Begenau, J., & Landvoigt, T. (2021). Financial regulation in a quantitative model of the modern banking system. *NBER Working Paper*, *w28501*. https://www.nber.org/papers/w28501

Begenau, J. (2020). Capital requirements, risk choice, and liquidity provision in a business-cycle model. *Journal of Financial Economics*, *136*(2), 355–378. https://doi.org/10.1016/J.JFINECO.2019.10.004

Begenau, J., Bigio, S., Majerovitz, J., & Vieyra, M. (2020). A q-theory of banks. https://doi.org/10.3386/W27935

Boissay, F., Collard, F., & Smets, F. (2016). Booms and banking crises. *Journal of Political Economy*, *124*(2), 489–538. https://doi.org/10.1086/685475

Brunnermeier, M. K., & Sannikov, Y. (2014). A macroeconomic model with a financial sector. *American Economic Review*, *104*(2), 379–421. https://doi.org/10.1257/aer.104.2.379

Buchak, G., Matvos, G., Piskorski, T., & Seru, A. (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics*, *130*(3), 453–483. https://doi.org/10.1016/J.JFINECO.2018.03.011

Canzoneri, M., Diba, B., Guerrieri, L., & Mishin, A. (2020). Optimal Dynamic Capital Requirements and Implementable Capital Buffer Rules. *Finance and Economics Discussion Series*, *2020*(056). https://doi.org/10.17016/feds.2020.056

Chernenko, S., & Sunderam, A. (2014). Frictions in Shadow Banking: Evidence from the Lending Behavior of Money Market Mutual Funds. *Review of Financial Studies*, *27*(6), 1717–1750. https://doi.org/10.1093/rfs/hhu025

Christiano, L. J., Eichenbaum, M., & Evans, C. L. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy*, *113*(1), 1–45. https://doi.org/10.1086/426038

Collard, F., Dellas, H., Diba, B., & Loisel, O. (2017). Optimal Monetary and Prudential Policies. *American Economic Journal: Macroeconomics*, *9*(1), 40–87. http://www.aeaweb.org/articles?id=10.1257/mac.20140139

Davydiuk, T. (2018). Dynamic Bank Capital Requirements. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3110800

Elenev, V., Landvoigt, T., & Nieuwerburgh, S. V. (2021). A Macroeconomic Model With Financially Constrained Producers and Intermediaries. *Econometrica*, *89*(3), 1361–1418. https://doi.org/10.3982/ECTA16438

Faria-e-Castro, M. (2020). A Quantitative Analysis of Countercyclical Capital Buffers. *SSRN*. https://doi.org/10.20955/WP.2019.008

Feenstra, R. (1986). Functional equivalence between liquidity costs and the utility of money. *Journal of Monetary Economics*, *17*, 271–291. https://doi.org/10.1016/0304-3932(86)90032-2

Ferrante, F. (2018). A Model of Endogenous Loan Quality and the Collapse of the Shadow Banking System. *American Economic Journal: Macroeconomics*, *10*(4), 152–201. https://doi.org/10.1257/mac.20160118

Financial Stability Report, F. E. D. (2019). The fed - financial stability report – may 2019 [Online; accessed 12-December-2022 via https://www.federalreserve.gov/publications/2019-may-financial-stability-report-purpose.htm].

Galí, J. (1999). Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *American Economic Review*, *89*(1), 249–271. https://doi.org/10.1257/AER.89.1.249

Gertler, M., Kiyotaki, N., & Prestipino, A. (2016). Wholesale Banking and Bank Runs in Macroeconomic Modeling of Financial Crises. *Handbook of Macroeconomics*, *2*, 1345–1425. https://doi.org/10.1016/BS.HESMAC.2016.03.009

Gertler, M., & Karadi, P. (2011). A model of unconventional monetary policy. *Journal of Monetary Economics*, *58*(1), 17–34. https://doi.org/10.1016/J.JMONECO.2010.10.004

Gertler, M., Kiyotaki, N., & Prestipino, A. (2019). A Macroeconomic Model with Financial Panics. *The Review of Economic Studies*. https://doi.org/10.1093/restud/rdz032

Gertler, M., Kiyotaki, N., & Prestipino, A. (2020). Credit booms, financial crises, and macroprudential policy. *Review of Economic Dynamics*, *37*, S8–S33. https://doi.org/10.1016/j.red.2020.06.004

Gilchrist, B. S., & Zakrajšek, E. (2012). Credit spreads and business cycle fluctuations. *American Economic Review*, *102*, 1692–1720. https://doi.org/10.1257/AER.102.4.1692

Gorton, G., Lewellen, S., & Metrick, A. (2012). The Safe-Asset Share. *American Economic Review*, *102*(3), 101–106. https://doi.org/10.1257/aer.102.3.101

Guerrieri, L., & Iacoviello, M. (2015). OccBin: A toolkit for solving dynamic models with occasionally binding constraints easily. *Journal of Monetary Economics*, *70*, 22–38. https://doi.org/10.1016/J.JMONECO.2014.08.005

He, Z., & Krishnamurthy, A. (2012). A model of capital and crises. *The Review of Economic Studies*, *79*(2), 735–777. https://doi.org/10.1093/restud/rdr036

He, Z., Khang, I. G., & Krishnamurthy, A. (2010). Balance Sheet Adjustments during the 2008 Crisis. *IMF Economic Review*, *58*(1), 118–156. https://doi.org/10.1057/imfer.2010.6

Jiang, E., Matvos, G., Piskorski, T., & Seru, A. (2020). Banking Without Deposits: Evidence from Shadow Bank Call Reports. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3584191

Krishnamurthy, A., & Vissing-Jorgensen, A. (2012). The Aggregate Demand for Treasury Debt. *Journal of Political Economy*, *120*(2), 233–267. https://doi.org/10.1086/666526

Malherbe, F. (2020). Optimal Capital Requirements over the Business and Financial Cycles. *American Economic Journal: Macroeconomics*, *12*(3), 139–74. https://doi.org/10.1257/MAC.20160140

Martinez-Miera, D., & Repullo, R. (2019). Markets, banks, and shadow banks. https://doi.org/10.2866/627556

Mendicino, C., Nikolov, K., Suarez, J., & Supera, D. (2018). Optimal Dynamic Capital Requirements. *Journal of Money, Credit and Banking*, *50*(6), 1271–1297. https://doi.org/10.1111/JMCB.12490

Mendoza, E. G. (2010). Sudden stops, financial crises, and leverage. *American Economic Review*, *100*(5), 1941–1966. https://doi.org/10.1257/aer.100.5.1941

Minneapolis Plan, F. E. D. (2017). The minneapolis plan to end too big to fail. minneapolis: Federal reserve bank of minneapolis [Online; accessed 12-December-2022 via https://www.minneapolisfed.org/~/media/files/publications/studies/endingtbtf/the-minneapolis-plan/the-minneapolis-plan-to-end-too-big-to-fail-final.pdf].

Nagel, S. (2016). The Liquidity Premium of Near-Money Assets. *The Quarterly Journal of Economics*, *131*(4), 1927–1971. https://doi.org/10.1093/qje/qjw028

Ordoñez, G. (2018). Sustainable Shadow Banking. *American Economic Journal: Macroeconomics*, *10*(1), 33–56. https://doi.org/10.1257/mac.20150346

Pozsar, Z., Adrian, T., Ashcraft, A. B., & Boesky, H. (2013). Shadow banking. *Economic Policy Review*, (Dec), 1–16. https://ideas.repec.org/a/fip/fednep/00001.html

Quadrini, V. (2017). Bank liabilities channel. *Journal of Monetary Economics*, *89*, 25–44. https://doi.org/10.1016/J.JMONECO.2017.03.006

Sunderam, A. (2015). Money Creation and the Shadow Banking System. *The Review of Financial Studies*, *28*(4), 939–977. https://doi.org/10.1093/rfs/hhu083

Van den Heuvel, S. J. (2008). The welfare cost of bank capital requirements. *Journal of Monetary Economics*, *55*(2), 298–320. https://doi.org/10.1016/j.jmoneco.2007.12.001

Zhang, J. (2020). Shadow banking and optimal capital requirements. *Review of Economic Dynamics*, *38*, 296–325. https://doi.org/10.1016/J.RED.2020.05.004

Online Appendix for
"Dynamic Bank Capital Regulation in the Presence of Shadow Banks"

November 2022

I omit the index and the type of bank in the expressions when it is evident from the context which bank I refer to.

A Expression of Net Cash Flow

Calculating the integral from the expression of net cash flows (suppressing the index i):

$$E_t\left\{\int_{\varepsilon^*_{t+1}}^{\infty}\left(\left(R^l_{t+1}+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}-R^d_{t+1}\right)l_t+R^d_{t+1}e_t\right)\,\mathrm{d}G(\varepsilon_{t+1})\right\},$$

where $\left(R^l_{t+1}+\sigma_t\frac{\varepsilon^*_{t+1}}{Q_t}-R^d_{t+1}\right)l_t+R^d_{t+1}e_t=0$.

Break calculation of the integral into two parts.

1. $\int_{\varepsilon^*_{t+1}}^{\infty}\varepsilon_{t+1}\,\mathrm{d}G(\varepsilon_{t+1})$,
2. $\int_{\varepsilon^*_{t+1}}^{\infty}\mathrm{d}G(\varepsilon_{t+1})$.

Working on the first part:

$$\int_{\varepsilon^*_{t+1}}^{\infty}\varepsilon_{t+1}\,\mathrm{d}G(\varepsilon_{t+1})=\int_{\left(\frac{R^d_{t+1}-R^l_{t+1}}{\sigma_t}-\frac{R^d_{t+1}e_t}{\sigma_t l_t}\right)Q_t}^{\infty}\varepsilon_{t+1}\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}=$$

Introduce a change in variables to recast the integral in terms of the Standard Normal distribution. Use $v=\frac{\varepsilon_{t+1}+\xi}{\sqrt{2}\tau}$, or equivalently $\varepsilon_{t+1}=v\sqrt{2}\tau-\xi$, and remember that for the change $x=\varphi(t)$, the integral$\int_{\varphi(a)}^{\varphi(b)}f(x)dx$ becomes $\int_a^b f(\varphi(t))\varphi'(t)dt$. Since $dv=\frac{d\varepsilon_{t+1}}{\sqrt{2}\tau}$, multiply dv by $\sqrt{2}\tau$ to express $d\varepsilon_{t+1}$ in terms of dv. Moreover, to transform the lower limit of the integral, add ξ to the current expression of the lower bound of the integral and divide the result by $\sqrt{2}\tau$.

$$\int\limits_{\frac{\left(R^d_{t+1}-R^l_{t+1}\right)l_tQ_t-R^d_{t+1}e_tQ_t+\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}}^{\infty}\left(v\sqrt{2}\tau-\xi\right)\frac{\sqrt{2}\tau}{\sqrt{2\pi\tau^2}}e^{-v^2}\,\mathrm{d}v=\int\limits_{\frac{\left(R^d_{t+1}-R^l_{t+1}\right)l_tQ_t-R^d_{t+1}e_tQ_t+\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}}^{\infty}\left(v\sqrt{2}\tau-\xi\right)\frac{1}{\sqrt{\pi}}e^{-v^2}\,\mathrm{d}v=$$

$$\frac{\sqrt{2}\tau}{\sqrt{\pi}}\int\limits_{\frac{\left(R^d_{t+1}-R^l_{t+1}\right)l_tQ_t-R^d_{t+1}e_tQ_t+\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}}^{\infty}ve^{-v^2}\,\mathrm{d}v-\frac{\xi}{\sqrt{\pi}}\int\limits_{\frac{\left(R^d_{t+1}-R^l_{t+1}\right)l_tQ_t-R^d_{t+1}e_tQ_t+\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}}^{\infty}e^{-v^2}\,\mathrm{d}v=$$

$$-\frac{\sqrt{2}\tau}{2\sqrt{\pi}}e^{-v^2}\Bigg|_{\frac{\left(R^d_{t+1}-R^l_{t+1}\right)l_tQ_t-R^d_{t+1}e_tQ_t+\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}}^{\infty}-\frac{\xi}{\sqrt{\pi}}\left[\int\limits_0^{\infty}e^{-v^2}\,\mathrm{d}v-\int\limits_0^{\frac{\left(R^d_{t+1}-R^l_{t+1}\right)l_tQ_t-R^d_{t+1}e_tQ_t+\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}}e^{-v^2}\,\mathrm{d}v\right]=$$

$$0+l_t\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R^l_{t+1}-R^d_{t+1}\right)l_tQ_t+R^d_{t+1}e_tQ_t-\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}\right)^2}-$$

$$\frac{\xi}{\sqrt{\pi}}\left[\frac{\sqrt{\pi}}{2}\mathrm{erf}(\infty)-\frac{\sqrt{\pi}}{2}\mathrm{erf}\left(\frac{\left(R^l_{t+1}-R^d_{t+1}\right)l_tQ_t+R^d_{t+1}e_tQ_t-\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}\right)\right]=$$

$$\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R^l_{t+1}-R^d_{t+1}\right)l_tQ_t+R^d_{t+1}e_tQ_t-\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}\right)^2}-\frac{\xi}{2}\left[1+\mathrm{erf}\left(\frac{\left(R^l_{t+1}-R^d_{t+1}\right)l_tQ_t+R^d_{t+1}e_tQ_t-\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}\right)\right],$$

where I use that $\mathrm{erf}(\infty)=1$ and $\mathrm{erf}(-x)=-\mathrm{erf}(x)$.

Working on the second part: Again, use the transformation $v=\frac{\varepsilon_{t+1}+\xi}{\sqrt{2}\tau}$, so $\varepsilon_{t+1}=v\sqrt{2}\tau-\xi$

$$\int\limits_{\varepsilon^*_{t+1}}^{\infty}\mathrm{d}G(\varepsilon_{t+1})=\int\limits_{\varepsilon^*_{t+1}}^{\infty}\left(\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\right)=$$

$$\int\limits_{\frac{\left(R^d_{t+1}-R^l_{t+1}\right)l_tQ_t-R^d_{t+1}e_tQ_t+\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}}^{\infty}\frac{\sqrt{2}\tau}{\sqrt{2\pi\tau^2}}e^{-v^2}\,\mathrm{d}v=\frac{1}{\sqrt{\pi}}\int\limits_{\frac{\left(R^d_{t+1}-R^l_{t+1}\right)l_tQ_t-R^d_{t+1}e_tQ_t+\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}}^{\infty}e^{-v^2}\,\mathrm{d}v=$$

$$\frac{1}{2}\left[1+\mathrm{erf}\left(\frac{\left(R^l_{t+1}-R^d_{t+1}\right)l_tQ_t+R^d_{t+1}e_tQ_t-\xi\sigma_tl_t}{\sigma_tl_t\sqrt{2}\tau}\right)\right].$$

Combining both parts, find

$$E_t\left\{\int\limits_{\varepsilon^*_{t+1}}^{\infty}\left(\left(R^l_{t+1}+\sigma_t\varepsilon_{t+1}-R^d_{t+1}\right)l_t+R^d_{t+1}e_t\right)\,\mathrm{d}G(\varepsilon_{t+1})\right\}=$$

$$E_t\left\{\sigma_t l_t\frac{\tau}{Q_t\sqrt{2\pi}}e^{-\left(\frac{\left(R^l_{t+1}-R^d_{t+1}\right)l_tQ_t+R^d_{t+1}e_tQ_t-\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)^2}+\right.$$

$$\left.\frac{\left(\left(R^l_{t+1}-R^d_{t+1}\right)l_t+R^d_{t+1}e_t-\frac{\xi}{Q_t}\right)}{2}\left[1+\operatorname{erf}\left(\frac{\left(R^l_{t+1}-R^d_{t+1}\right)l_tQ_t+R^d_{t+1}e_tQ_t-\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right]\right\}.$$

B Share of Non-Defaulted Deposits

First, the bank defaults on its deposit obligation whenever the idiosyncratic shock, ε_{t+1}, is below a cutoff level ε^*_{t+1}, defined by:

$$\varepsilon^*_t=-\frac{\left(R^l_t-R^d_t\right)l_{t-1}Q_{t-1}+R^d_te_{t-1}Q_{t-1}}{\sigma_{t-1}l_{t-1}}.$$

Second, let's express $1+\operatorname{erf}\left(\frac{\left(R^l_t-R^d_t\right)l_{t-1}Q_{t-1}+R^d_te_{t-1}Q_{t-1}-\xi\sigma_{t-1}l_{t-1}}{\sigma_{t-1}l_{t-1}\sqrt{2}\tau}\right)=1-\operatorname{erf}\left(\frac{\varepsilon^*_t+\xi}{\tau\sqrt{2}}\right)$ in terms of the CDF of the normal distribution. By definition,

$$G(\varepsilon_t)=\frac{1}{2}\left[1+\operatorname{erf}\left(\frac{\varepsilon_t+\xi}{\tau\sqrt{2}}\right)\right]\qquad\Longrightarrow\qquad\operatorname{erf}\left(\frac{\varepsilon_t+\xi}{\tau\sqrt{2}}\right)=2G(\varepsilon_t)-1.$$

Therefore,

$$1+\operatorname{erf}\left(\frac{\left(R^l_t-R^d_t\right)l_{t-1}Q_{t-1}+R^d_te_{t-1}Q_{t-1}-\xi\sigma_{t-1}l_{t-1}}{\sigma_{t-1}l_{t-1}\sqrt{2}\tau}\right)=1-\left(2G(\varepsilon^*_t)-1\right)=2\left(1-G(\varepsilon^*_t)\right).$$

Thus, the share of defaulted loans is given by

$$1-G(\varepsilon^*_t)=\frac{1}{2}\left(1+\operatorname{erf}\left(\frac{\left(R^l_t-R^d_t\right)l_{t-1}Q_{t-1}+R^d_te_{t-1}Q_{t-1}-\xi\sigma_{t-1}l_{t-1}}{\sigma_{t-1}l_{t-1}\sqrt{2}\tau}\right)\right)$$

C Choice of Risk

Theorem. *The expected dividends function of banks, $E_t\omega_{t+1}$, is convex in the risk parameter σ_t. Moreover, if $\xi = 0$, then it is also increasing in σ_t.*

Proof. I generalize the proof taken from Van den Heuvel (2008) to the case with aggregate uncertainty. The proof applies to an arbitrary distribution of the idiosyncratic shock, ε_{t+1}, so a Normal distribution considered in the analysis is not a special case chosen to drive the results. $\square$

Assumption. *ε has a cumulative distribution function G_ε with support $[\underline{\varepsilon},\ \bar{\varepsilon}]$, with $\underline{\varepsilon} < 0 < \bar{\varepsilon}$. The mean of ε is equal to $-\xi$. ε is independent of the aggregate shock. The aggregate shock does not depend on the choice of σ_t.*

Note that I do not restrict the analysis to the bounded support[36], so $\underline{\varepsilon}$ and $\bar{\varepsilon}$ can take $-\infty$ and $+\infty$, respectively. Note that G_ε need not be continuous.

Let $\hat{\varepsilon}(\sigma_t,\ R^l_{t+1}) \equiv \frac{R^d_t d_t Q_t}{\sigma_t l_t} - \frac{R^l_{t+1} Q_t}{\sigma_t}$, so $\left(R^l_{t+1} + \sigma_t \frac{\hat{\varepsilon}(\sigma_t)}{Q_t}\right) l_t - R^d_t d_t = 0$.

Let $\pi(\sigma_t,\ R^l_{t+1}) = E_\varepsilon \left[\left(\left(R^l_{t+1} + \sigma_t \frac{\varepsilon}{Q_t}\right) l_t - R^d_t d_t\right)^+\right]$ be a function of expected dividends (taken over the idiosyncratic shock only) under some realization of R^l_{t+1} which is considered to be fixed in this function. To account for the aggregate uncertainty, R^l_{t+1} needs to be a random variable. Therefore, the expected dividends are given by (taking into account both idiosyncratic and aggregate) uncertainty:

$$\Pi(\sigma_t) = \int_\Omega \pi\left(\sigma_t,\ R^l_{t+1}(\omega)\right) P(d\omega) = E_t \left[\int_{\hat{\varepsilon}(\sigma_t,\ R^l_{t+1})}^{\bar{\varepsilon}} \left(\left(R^l_{t+1} + \sigma_t \frac{\varepsilon}{Q_t}\right) l_t - R^d_t d_t\right) dG_\varepsilon\right] =$$

$$E_t \left[\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left(\left(R^l_{t+1} + \sigma_t \frac{\varepsilon}{Q_t}\right) l_t - R^d_t d_t\right) dG_\varepsilon\right] - E_t \left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t,\ R^l_{t+1})} \left(\left(R^l_{t+1} + \sigma_t \frac{\varepsilon}{Q_t}\right) l_t - R^d_t d_t\right) dG_\varepsilon\right] =$$

$$E_t R^l_{t+1} l_t - R^d_t d_t - \sigma_t \frac{\xi}{Q_t} l_t - \frac{\sigma_t l_t}{Q_t} E_t \left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t,\ R^l_{t+1})} \left(\varepsilon - \hat{\varepsilon}(\sigma_t,\ R^l_{t+1})\right) dG_\varepsilon\right] =$$

$$E_t R^l_{t+1} l_t - R^d_t d_t + \frac{l_t}{Q_t} \left(\sigma_t E_t \left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t,\ R^l_{t+1})} \left(\hat{\varepsilon}(\sigma_t,\ R^l_{t+1}) - \varepsilon\right) dG_\varepsilon\right] - \sigma_t \xi\right).$$

Note that $\left(R^l_{t+1} + \sigma_t \frac{\varepsilon}{Q_t}\right) l_t - R^d_t d_t$ is expressed in terms of $\hat{\varepsilon}(\sigma_t,\ R^l_{t+1})$ and ε using the definition of $\hat{\varepsilon}(\sigma_t,\ R^l_{t+1})$.

The proof below shows that $\Pi(\sigma_t)$ is convex in σ_t. Since the expression of $\Pi(\sigma_t)$ involves the term which is linear in σ_t and $\frac{l_t}{Q_t} \geq 0$, the sufficient condition for $\Pi(\sigma_t)$ to be convex in

[36]Unbounded support is more relevant for aggregate shocks

σ_t is that

$$H(\sigma_t) \equiv E_t \left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t)} \left(\hat{\varepsilon}(\sigma_t) - \varepsilon \right) dG_\varepsilon \right] \sigma_t$$

is convex in σ_t,.

Claim. $H(\sigma_t) \equiv E_t \left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t)} \left(\hat{\varepsilon}(\sigma_t,\, R^l_{t+1}) - \varepsilon \right) dG_\varepsilon \right] \sigma_t$ *is convex* σ_t*. It is also increasing in* σ_t *when* $\xi = 0$*.*

Proof. Steps of the proof: □

1. Define $h(\sigma_t,\, R^l_{t+1}) \equiv \sigma_t \left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t,\, R^l_{t+1})} \left(\hat{\varepsilon}(\sigma_t,\, R^l_{t+1}) - \varepsilon \right) dG_\varepsilon \right]$ and consider 3 cases:
   (a) Realization of R^l_{t+1} is such that $\hat{\varepsilon}\left(\sigma_t,\, R^l_{t+1}\right) = \frac{R^d_t d_t Q_t}{\sigma_t l_t} - \frac{R^l_{t+1} Q_t}{\sigma_t} > 0$,
   (b) Realization of R^l_{t+1} is such that $\hat{\varepsilon}\left(\sigma_t,\, R^l_{t+1}\right) = \frac{R^d_t d_t Q_t}{\sigma_t l_t} - \frac{R^l_{t+1} Q_t}{\sigma_t} < 0$,
   (c) Realization of R^l_{t+1} is such that $\hat{\varepsilon}\left(\sigma_t,\, R^l_{t+1}\right) = \frac{R^d_t d_t Q_t}{\sigma_t l_t} - \frac{R^l_{t+1} Q_t}{\sigma_t} = 0$,

   Show that $h(\sigma_t,\, R^l_{t+1})$ is convex and increasing in σ_t in cases 1a and 1b and $h(\sigma_t,\, R^l_{t+1})$ is linear and increasing in σ_t in case 1c.

2. Employ the argument that convexity and monotonicity are preserved under non-negative scaling and addition (guaranteed by the expectation operator over the aggregate uncertainty) to find that $H(\sigma_t)$ is convex and increasing.

Here I show each step of the proof formally

1. Let $\sigma_{1t} < \sigma_{2t}$ and, for $\lambda \in (0,\, 1)$, define $\sigma_{\lambda t} = \lambda \sigma_{1t} + (1 - \lambda)\sigma_{2t}$. Let $\hat{\varepsilon}_i = \hat{\varepsilon}(\sigma_{it},\, R^l_{t+1}) \equiv \frac{R^d_t d_t Q_t}{\sigma_t l_t} - \frac{R^l_{t+1} Q_t}{\sigma_t}$, for $i = 1,\, 2,\, \lambda$.
   (a) $\hat{\varepsilon}\left(\sigma_t,\, R^l_{t+1}\right) > 0$ implies that $0 < \hat{\varepsilon}_2 < \hat{\varepsilon}_\lambda < \hat{\varepsilon}_1$,

Claim. $h(\sigma_t)$ is convex in σ_t.

$$h(\sigma_{\lambda t}) = (\lambda\sigma_{1t} + (1-\lambda)\sigma_{2t}) \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_{\lambda t})} (\hat{\varepsilon}(\sigma_{\lambda t}) - \varepsilon)\, dG_\varepsilon \right\} =$$

$$\lambda\sigma_{1t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon - \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \right\} +$$

$$(1-\lambda)\sigma_{2t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon + \int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \right\} =$$

$$\lambda\sigma_{1t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_1) + \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\varepsilon - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \right\} +$$

$$(1-\lambda)\sigma_{2t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_2) + \int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \right\} \leq$$

$$\lambda\sigma_{1t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_1) + \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \right\} +$$

$$(1-\lambda)\sigma_{2t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_2) + \int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, dG_\varepsilon \right\},$$

as $\int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\varepsilon - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \leq \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \hat{\varepsilon}_\lambda)\, dG_\varepsilon$ and $\int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \leq \int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, dG_\varepsilon$. Using $h(\sigma_{1t}) = \sigma_{1t} \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon$ and $h(\sigma_{2t}) = \sigma_{2t} \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon$, one can get:

$$h(\sigma_{\lambda t}) \leq \lambda h(\sigma_{1t}) + (1-\lambda)h(\sigma_{2t}) + \lambda\sigma_{1t} \left\{ (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_\lambda) \right\} + (1-\lambda)\sigma_{2t} \left\{ (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_\lambda) \right\} =$$

$$\lambda h(\sigma_{1t}) + (1-\lambda)h(\sigma_{2t}) + G_\varepsilon(\hat{\varepsilon}_\lambda) \left(\lambda\sigma_{1t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1) + (1-\lambda)\sigma_{2t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2) \right) =$$

$$\lambda h(\sigma_{1t}) + (1-\lambda)h(\sigma_{2t}),$$

where the last equality follows from $\sigma_{1t}\hat{\varepsilon}_1 = \frac{R_t^d d_t Q_t}{l_t} - R_{t+1}^l Q_t = \sigma_{2t}\hat{\varepsilon}_2 = \sigma_{\lambda t}\hat{\varepsilon}_\lambda$. So

$$\lambda\sigma_{1t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1) + (1-\lambda)\sigma_{2t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2) =$$

$$\hat{\varepsilon}_\lambda (\lambda\sigma_{1t} + (1-\lambda)\sigma_{2t}) - \left(\frac{R_t^d d_t Q_t}{l_t} - R_{t+1}^l Q_t \right) (\lambda + (1-\lambda)) =$$

$$\sigma_{\lambda t}\hat{\varepsilon}_\lambda - \left(\frac{R_t^d d_t Q_t}{l_t} - R_{t+1}^l Q_t \right) = 0.$$

Claim. If $\xi = 0$, then $h(\sigma_t)$ is increasing in σ_t.

$$h(\sigma_{2t}) - h(\sigma_{1t}) = \sigma_{2t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon - \sigma_{1t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon =$$

$$\sigma_{2t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon - \sigma_{1t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon - \sigma_{1t}\int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon =$$

$$\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2\sigma_{2t} - \varepsilon\sigma_{2t} - \hat{\varepsilon}_1\sigma_{1t} + \varepsilon\sigma_{1t})\, dG_\varepsilon - \sigma_{1t}\int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon =$$

$$-\sigma_{2t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} \varepsilon dG_\varepsilon - \sigma_{1t}\left(\int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon - \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} \varepsilon dG_\varepsilon\right) =$$

$$-\sigma_{2t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} \varepsilon dG_\varepsilon - \sigma_{1t}\left(\int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_1} \hat{\varepsilon}_1 dG_\varepsilon - \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} \varepsilon dG_\varepsilon\right) =$$

$$-\sigma_{2t}\left(0 - \int_{\hat{\varepsilon}_2}^{\bar{\varepsilon}} \varepsilon dG_\varepsilon\right) - \sigma_{1t}\int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_1} \hat{\varepsilon}_1 dG_\varepsilon + \sigma_{1t}\left(0 - \int_{\hat{\varepsilon}_1}^{\bar{\varepsilon}} \varepsilon dG_\varepsilon\right) =$$

$$\sigma_{2t}\int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_1} \varepsilon dG_\varepsilon + \sigma_{2t}\int_{\hat{\varepsilon}_1}^{\bar{\varepsilon}} \varepsilon dG_\varepsilon - \sigma_{1t}\int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_1} \hat{\varepsilon}_1 dG_\varepsilon - \sigma_{1t}\int_{\hat{\varepsilon}_1}^{\bar{\varepsilon}} \varepsilon dG_\varepsilon >$$

$$(\sigma_{2t} - \sigma_{1t})\int_{\hat{\varepsilon}_1}^{\bar{\varepsilon}} \varepsilon dG_\varepsilon + \sigma_{2t}\int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_1} \hat{\varepsilon}_2 dG_\varepsilon - \sigma_{1t}\int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_1} \hat{\varepsilon}_1 dG_\varepsilon = (\sigma_{2t} - \sigma_{1t})\int_{\hat{\varepsilon}_1}^{\bar{\varepsilon}} \varepsilon dG_\varepsilon > 0$$

(b) $\hat{\varepsilon}\left(\sigma_t,\, R^l_{t+1}\right) < 0$ implies that $\hat{\varepsilon}_1 < \hat{\varepsilon}_\lambda < \hat{\varepsilon}_2 < 0$

Claim. $h(\sigma_t)$ is convex in σ_t.

$$h(\sigma_{\lambda t}) = (\lambda\sigma_{1t} + (1-\lambda)\sigma_{2t})\left\{\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_{\lambda t})} (\hat{\varepsilon}(\sigma_{\lambda t}) - \varepsilon)\, dG_\varepsilon\right\} =$$

$$\lambda\sigma_{1t}\left\{\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon + \int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon\right\} +$$

$$(1-\lambda)\sigma_{2t}\left\{\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon - \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon\right\} =$$

$$\lambda\sigma_{1t}\left\{\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_1) + \int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon\right\} +$$

$$(1-\lambda)\sigma_{2t}\left\{\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_2) + \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\varepsilon - \hat{\varepsilon}_\lambda)\, dG_\varepsilon\right\} \leq$$

$$\lambda\sigma_{1t}\left\{\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_1) + \int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, dG_\varepsilon\right\} +$$

$$(1-\lambda)\sigma_{2t}\left\{\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_2) + \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \hat{\varepsilon}_\lambda)\, dG_\varepsilon\right\},$$

as $\int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda}(\hat{\varepsilon}_\lambda-\varepsilon)\,dG_\varepsilon \le \int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda}(\hat{\varepsilon}_\lambda-\hat{\varepsilon}_1)\,dG_\varepsilon$ and $\int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2}(\varepsilon-\hat{\varepsilon}_\lambda)\,dG_\varepsilon \le \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2}(\hat{\varepsilon}_2-\hat{\varepsilon}_\lambda)\,dG_\varepsilon$. Using $h(\sigma_{1t})=\sigma_{1t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1}(\hat{\varepsilon}_1-\varepsilon)\,dG_\varepsilon$ and $h(\sigma_{2t})=\sigma_{2t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2}(\hat{\varepsilon}_2-\varepsilon)\,dG_\varepsilon$, one can get:

$$
\begin{aligned}
h(\sigma_{\lambda t}) \le \lambda h(\sigma_{1t})+(1-\lambda)h(\sigma_{2t})+\lambda\sigma_{1t}\left\{(\hat{\varepsilon}_\lambda-\hat{\varepsilon}_1)\,G_\varepsilon(\hat{\varepsilon}_\lambda)\right\}+(1-\lambda)\sigma_{2t}\left\{(\hat{\varepsilon}_\lambda-\hat{\varepsilon}_2)\,G_\varepsilon(\hat{\varepsilon}_\lambda)\right\} = \\
\lambda h(\sigma_{1t})+(1-\lambda)h(\sigma_{2t})+G_\varepsilon(\hat{\varepsilon}_\lambda)\left(\lambda\sigma_{1t}(\hat{\varepsilon}_\lambda-\hat{\varepsilon}_1)+(1-\lambda)\sigma_{2t}(\hat{\varepsilon}_\lambda-\hat{\varepsilon}_2)\right) = \\
\lambda h(\sigma_{1t})+(1-\lambda)h(\sigma_{2t}),
\end{aligned}
$$

where the last equality applies the similar reasoning used for the previous case. Therefore, $h(\sigma_t)$ is convex in σ_t for $R^l_{t+1} > R^d_t(1-\gamma)$.

Claim. $h(\sigma_t)$ is increasing in σ_t.

$$
\begin{aligned}
h(\sigma_{2t})-h(\sigma_{1t}) = \sigma_{2t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2}(\hat{\varepsilon}_2-\varepsilon)\,dG_\varepsilon-\sigma_{1t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1}(\hat{\varepsilon}_1-\varepsilon)\,dG_\varepsilon = \\
\sigma_{2t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1}(\hat{\varepsilon}_2-\varepsilon)\,dG_\varepsilon+\sigma_{2t}\int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_2}(\hat{\varepsilon}_2-\varepsilon)\,dG_\varepsilon-\sigma_{1t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1}(\hat{\varepsilon}_1-\varepsilon)\,dG_\varepsilon = \\
\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1}(\hat{\varepsilon}_2\sigma_{2t}-\varepsilon\sigma_{2t}-\hat{\varepsilon}_1\sigma_{1t}+\varepsilon\sigma_{1t})\,dG_\varepsilon+\sigma_{2t}\int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_2}(\hat{\varepsilon}_2-\varepsilon)\,dG_\varepsilon = \\
\sigma_{1t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1}\varepsilon dG_\varepsilon+\sigma_{2t}\left(\int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_2}(\hat{\varepsilon}_2-\varepsilon)\,dG_\varepsilon-\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1}\varepsilon dG_\varepsilon\right) = \\
\sigma_{1t}\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1}\varepsilon dG_\varepsilon+\sigma_{2t}\left(\int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_2}\hat{\varepsilon}_2 dG_\varepsilon-\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2}\varepsilon dG_\varepsilon\right) = \\
(\sigma_{1t}-\sigma_{2t})\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1}\varepsilon dG_\varepsilon+\sigma_{2t}\left(\int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_2}\hat{\varepsilon}_2 dG_\varepsilon-\int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_2}\varepsilon dG_\varepsilon\right) = \\
(\sigma_{1t}-\sigma_{2t})\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1}\varepsilon dG_\varepsilon+\sigma_{2t}\left(\int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_2}(\hat{\varepsilon}_2-\varepsilon)\,dG_\varepsilon\right) > 0
\end{aligned}
$$

because $\sigma_{1t}-\sigma_{2t}<0$, $\int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1}\varepsilon dG_\varepsilon<0$ and $\int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_2}(\hat{\varepsilon}_2-\varepsilon)\,dG_\varepsilon>0$

(c) $\hat{\varepsilon}\left(\sigma_t,\, R^l_{t+1}\right)=0$ implies $h(\sigma_t)=\sigma_t\left[\int_{\underline{\varepsilon}}^{0}(0-\varepsilon)\,dG_\varepsilon\right]$ is linear and increasing in σ_t.

2. Consider $P(\omega)\ge 0$ for each $R^l_{t+1}(\omega)\in R$. Then the following function[37]:

$$
\int_\Omega h\left(\sigma_t,\, R^l_{t+1}(\omega)\right)P(d\omega)=E_t\,h(\sigma_t,\, R^l_{t+1})\equiv H(\sigma_t)
$$

is convex in σ_t and increasing in σ_t when $\xi=0$. It follows directly from the linearity of the expectation operator which puts a non-negative weight on every realization of R^l_{t+1}

[37] Linearity in σ_t for one particular value of R^l_{t+1} can be considered as a weakly convex function, so it does not change the nature of the argument

and the fact that the sum of convex functions is a convex function. Therefore,

$$\Pi(\sigma_t) = R^l_{t+1} l_t - R^d_t d_t + \frac{l_t}{Q_t} H(\sigma_t)$$

is convex and also increasing in σ_t when $\xi = 0$ as $\frac{l_t}{Q_t} > 0$. $\square$

D The Shadow Bank's Problem

D.1 Value function

Substituting $\xi = 0$ into the expression of net cash flow to write the value of the bank:

$$V^S_{j,t} = E_t \left\{ \sum_{i=0}^{\infty} (1-\theta)\,\theta^i \Lambda_{t,t+1+i} \left[\sigma^S_{j,t+i} l^S_{j,t+i} \frac{\tau}{Q_{t+i}\sqrt{2\pi}} e^{-\left(\frac{\left(R^l_{t+1+i} - R^{dS}_{t+1+i}\right) Q_{t+i}}{\sigma^S_{j,t+i}\sqrt{2}\tau} + \frac{R^{dS}_{t+1+i} e^S_{j,t+i} Q_{t+i}}{\sigma^S_{j,t+i} l^S_{j,t+i}\sqrt{2}\tau} \right)^2} + \right.\right.$$
$$\left.\left. \frac{\left(\left(R^l_{t+1+i} - R^{dS}_{t+1+i}\right) l^S_{j,t+i} + R^{dS}_{t+1+i} e^S_{j,t+i}\right)}{2} \left[1 + \operatorname{erf}\left(\frac{\left(R^l_{t+1+i} - R^{dS}_{t+1+i}\right) Q_{t+i}}{\sigma^S_{j,t+i}\sqrt{2}\tau} + \frac{R^{dS}_{t+1+i} e^S_{j,t+i} Q_{t+i}}{\sigma^S_{j,t+i} l^S_{j,t+i}\sqrt{2}\tau} \right) \right] \right] \right\}.$$

Define

$$\nu_{j,t} = E_t \left\{ \sum_{i=0}^{\infty} (1-\theta)\,\theta^i \Lambda_{t,t+1+i} \left[\left(\sigma^S_{j,t+i} \frac{\tau}{Q_{t+i}\sqrt{2\pi}} e^{-\left(\frac{\left(R^l_{t+1+i} - R^{dS}_{t+1+i}\right) Q_{t+i}}{\sigma^S_{j,t+i}\sqrt{2}\tau} + \frac{R^{dS}_{t+1+i} e^S_{j,t+i} Q_{t+i}}{\sigma^S_{j,t+i} l^S_{j,t+i}\sqrt{2}\tau} \right)^2} + \right.\right.\right.$$
$$\left.\left.\left. \frac{\left(R^l_{t+1+i} - R^{dS}_{t+1+i}\right)}{2} \left[1 + \operatorname{erf}\left(\frac{\left(R^l_{t+1+i} - R^{dS}_{t+1+i}\right) Q_{t+i}}{\sigma^S_{j,t+i}\sqrt{2}\tau} + \frac{R^{dS}_{t+1+i} e^S_{j,t+i} Q_{t+i}}{\sigma^S_{j,t+i} l^S_{j,t+i}\sqrt{2}\tau} \right) \right] \right) \frac{l^S_{j,t+i}}{l^S_{j,t}} \right] \right\},$$

$$\eta_{j,t} = E_t \sum_{i=0}^{\infty} (1-\theta)\,\theta^i \Lambda_{t,t+1+i} \left[\frac{\left(R^{dS}_{t+1+i}\right)}{2} \left[1 + \operatorname{erf}\left(\frac{\left(R^l_{t+1+i} - R^{dS}_{t+1+i}\right) Q_{t+i}}{\sigma^S_{j,t+i}\sqrt{2}\tau} + \frac{R^{dS}_{t+1+i} e^S_{j,t+i} Q_{t+i}}{\sigma^S_{j,t+i} l^S_{j,t+i}\sqrt{2}\tau} \right) \right] \frac{e^S_{j,t+i}}{e^S_{j,t}} \right].$$

So $V_{j,t}^{S}=\upsilon_{j,t}l_{j,t}^{S}+\eta_{j,t}e_{j,t}^{S}$. Write it recursively. Pulling out the first term in each summation:

$$\nu_{j,t}=E_t\left\{(1-\theta)\Lambda_{t,t+1}\left[\left(\sigma_{j,t}^{S}\frac{\tau}{Q_t\sqrt{2\pi}}e^{-\left(\frac{\left(R_{t+1}^{l}-R_{t+1}^{dS}\right)Q_t}{\sigma_{j,t}^{S}\sqrt{2}\tau}+\frac{R_{t+1}^{dS}e_{j,t}^{S}Q_t}{\sigma_{j,t}^{S}l_{j,t}^{S}\sqrt{2}\tau}\right)^2}+\right.\right.\right.$$
$$\left.\left.\left.\frac{\left(R_{t+1}^{l}-R_{t+1}^{dS}\right)}{2}\left[1+\operatorname{erf}\left(\frac{\left(R_{t+1}^{l}-R_{t+1}^{dS}\right)Q_t}{\sigma_{j,t}^{S}\sqrt{2}\tau}+\frac{R_{t+1}^{dS}e_{j,t}^{S}Q_t}{\sigma_{j,t}^{S}l_{j,t}^{S}\sqrt{2}\tau}\right)\right]\right)\frac{l_{j,t}^{S}}{l_{j,t}^{S}}\right]\right\}+$$
$$E_t\left\{\sum_{i=1}^{\infty}(1-\theta)\theta^{i}\Lambda_{t,t+1+i}\left[\left(\sigma_{j,t+i}^{S}\frac{\tau}{Q_{t+i}\sqrt{2\pi}}e^{-\left(\frac{\left(R_{t+1+i}^{l}-R_{t+1+i}^{dS}\right)Q_{t+i}}{\sigma_{j,t+i}^{S}\sqrt{2}\tau}+\frac{R_{t+1+i}^{dS}e_{j,t+i}^{S}Q_{t+i}}{\sigma_{j,t+i}^{S}l_{j,t+i}^{S}\sqrt{2}\tau}\right)^2}+\right.\right.\right.$$
$$\left.\left.\left.\frac{\left(R_{t+1+i}^{l}-R_{t+1+i}^{dS}\right)}{2}\left[1+\operatorname{erf}\left(\frac{\left(R_{t+1+i}^{l}-R_{t+1+i}^{dS}\right)Q_{t+i}}{\sigma_{j,t+i}^{S}\sqrt{2}\tau}+\frac{R_{t+1+i}^{dS}e_{j,t+i}^{S}Q_{t+i}}{\sigma_{j,t+i}^{S}l_{j,t+i}^{S}\sqrt{2}\tau}\right)\right]\right)\frac{l_{j,t+i}^{S}}{l_{j,t}^{S}}\right]\right\}.$$

Transform the summations, so that they start from zero

$$\nu_{j,t}=E_t\left\{(1-\theta)\Lambda_{t,t+1}\left[\left(\sigma_{j,t}^{S}\frac{\tau}{Q_t\sqrt{2\pi}}e^{-\left(\frac{\left(R_{t+1}^{l}-R_{t+1}^{dS}\right)Q_t}{\sigma_{j,t}^{S}\sqrt{2}\tau}+\frac{R_{t+1}^{dS}e_{j,t}^{S}Q_t}{\sigma_{j,t}^{S}l_{j,t}^{S}\sqrt{2}\tau}\right)^2}+\right.\right.\right.$$
$$\left.\left.\left.\frac{\left(R_{t+1}^{l}-R_{t+1}^{dS}\right)}{2}\left[1+\operatorname{erf}\left(\frac{\left(R_{t+1}^{l}-R_{t+1}^{dS}\right)Q_t}{\sigma_{j,t}^{S}\sqrt{2}\tau}+\frac{R_{t+1}^{dS}e_{j,t}^{S}Q_t}{\sigma_{j,t}^{S}l_{j,t}^{S}\sqrt{2}\tau}\right)\right]\right)\right]\right\}+$$
$$E_t\left\{\theta\sum_{i=0}^{\infty}(1-\theta)\theta^{i}\Lambda_{t,t+2+i}\left[\left(\sigma_{j,t+1+i}^{S}\frac{\tau}{Q_{t+1+i}\sqrt{2\pi}}e^{-\left(\frac{\left(R_{t+2+i}^{l}-R_{t+2+i}^{dS}\right)Q_{t+1+i}}{\sigma_{j,t+1+i}^{S}\sqrt{2}\tau}+\frac{R_{t+2+i}^{dS}e_{j,t+1+i}^{S}Q_{t+1+i}}{\sigma_{j,t+1+i}^{S}l_{j,t+1+i}^{S}\sqrt{2}\tau}\right)^2}+\right.\right.\right.$$
$$\left.\left.\left.\frac{\left(\left(R_{t+2+i}^{l}-R_{t+2+i}^{dS}\right)\right)}{2}\left[1+\operatorname{erf}\left(\frac{\left(R_{t+2+i}^{l}-R_{t+2+i}^{dS}\right)Q_{t+1+i}}{\sigma_{j,t+1+i}^{S}\sqrt{2}\tau}+\frac{R_{t+2+i}^{dS}e_{j,t+1+i}^{S}Q_{t+1+i}}{\sigma_{j,t+1+i}^{S}l_{j,t+1+i}^{S}\sqrt{2}\tau}\right)\right]\right)\frac{l_{j,t+1+i}^{S}}{l_{j,t}^{S}}\right]\right\}.$$

Remember that $\Lambda_{t,t+2+i}=\Lambda_{t,t+1}\Lambda_{t+1,t+2+i}$. Plugging this into the last expressions

$$\nu_{j,t}=E_t\left\{(1-\theta)\Lambda_{t,t+1}\left[\left(\sigma_{j,t}^{S}\frac{\tau}{Q_t\sqrt{2\pi}}e^{-\left(\frac{\left(R_{t+1}^{l}-R_{t+1}^{dS}\right)Q_t}{\sigma_{j,t}^{S}\sqrt{2}\tau}+\frac{R_{t+1}^{dS}e_{j,t}^{S}Q_t}{\sigma_{j,t}^{S}l_{j,t}^{S}\sqrt{2}\tau}\right)^2}+\right.\right.\right.$$
$$\left.\left.\left.\frac{\left(R_{t+1}^{l}-R_{t+1}^{dS}\right)}{2}\left[1+\operatorname{erf}\left(\frac{\left(R_{t+1}^{l}-R_{t+1}^{dS}\right)Q_t}{\sigma_{j,t}^{S}\sqrt{2}\tau}+\frac{R_{t+1}^{dS}e_{j,t}^{S}Q_t}{\sigma_{j,t}^{S}l_{j,t}^{S}\sqrt{2}\tau}\right)\right]\right)\right]\right\}+$$

$$E_t\left\{\theta\Lambda_{t,t+1}\frac{l^S_{j,t+1}}{l^S_{j,t}}\sum_{i=0}^{\infty}\left(1-\theta\right)\theta^i\Lambda_{t+1,t+2+i}\left[\left(\left(\frac{\sigma^S_{j,t+1+i}}{Q_{t+1+i}}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R^l_{t+2+i}-R^{dS}_{t+2+i}\right)Q_{t+1+i}}{\sigma^S_{j,t+1+i}\sqrt{2}\tau}+\frac{R^{dS}_{t+2+i}e^S_{j,t+1+i}Q_{t+1+i}}{\sigma^S_{j,t+1+i}l^S_{j,t+1+i}\sqrt{2}\tau}\right)^2}+\right.\right.\right.\right.$$
$$\left.\left.\left.\left.\frac{\left(\left(R^l_{t+2+i}-R^{dS}_{t+2+i}\right)\right)}{2}\left[1+\text{erf}\left(\frac{\left(R^l_{t+2+i}-R^{dS}_{t+2+i}\right)Q_{t+1+i}}{\sigma^S_{j,t+1+i}\sqrt{2}\tau}+\frac{R^{dS}_{t+2+i}e^S_{j,t+1+i}Q_{t+1+i}}{\sigma^S_{j,t+1+i}l^S_{j,t+1+i}\sqrt{2}\tau}\right)\right]\right)\frac{l^S_{j,t+1+i}}{l^S_{j,t+1}}\right]\right\}.$$

Therefore,

$$\begin{aligned}\nu_{j,t}=&E_t\left\{\left(1-\theta\right)\Lambda_{t,t+1}\left[\sigma^S_{j,t}\frac{\tau}{Q_t\sqrt{2\pi}}e^{-\left(\frac{\left(R^l_{t+1}-R^{dS}_{t+1}\right)Q_t}{\sigma^S_{j,t}\sqrt{2}\tau}+\frac{R^{dS}_{t+1}e^S_{j,t}Q_t}{\sigma^S_{j,t}l^S_{j,t}\sqrt{2}\tau}\right)^2}+\right.\right.\\ &\left.\left.\frac{\left(R^l_{t+1}-R^{dS}_{t+1}\right)}{2}\left[1+\text{erf}\left(\frac{\left(R^l_{t+1}-R^{dS}_{t+1}\right)Q_t}{\sigma^S_{j,t}\sqrt{2}\tau}+\frac{R^{dS}_{t+1}e^S_{j,t}Q_t}{\sigma^S_{j,t}l^S_{j,t}\sqrt{2}\tau}\right)\right]\right]+\theta\Lambda_{t,t+1}\frac{l^S_{j,t+1}}{l^S_{j,t}}\nu_{j,t+1}\right\},\\ \eta_{j,t}=&E_t\left\{\left(1-\theta\right)\Lambda_{t,t+1}\left[\frac{\left(R^{dS}_{t+1}\right)}{2}\left[1+\text{erf}\left(\frac{\left(R^l_{t+1}-R^{dS}_{t+1}\right)Q_t}{\sigma^S_{j,t}\sqrt{2}\tau}+\frac{R^{dS}_{t+1}e^S_{j,t}Q_t}{\sigma^S_{j,t}l^S_{j,t}\sqrt{2}\tau}\right)\right]+\theta\Lambda_{t,t+1}\frac{e^S_{j,t+1}}{e^S_{j,t}}\eta_{j,t+1}\right]\right\}.\end{aligned}$$

D.2 Risk Choice

Denoting $\Gamma^S_{j,t}$ as expected dividends:

$$\Gamma^S_{j,t}=E_t\left\{\left(1-\theta\right)\Lambda_{t,t+1}\left[\left(\sigma^S_{j,t}l^S_{j,t}\frac{\tau}{Q_t\sqrt{2\pi}}e^{-\left(\frac{\left(R^l_{t+1}-R^{dS}_{t+1}\right)Q_t}{\sigma^S_{j,t}\sqrt{2}\tau}+\frac{R^{dS}_{t+1}e^S_{j,t}Q_t}{\sigma^S_{j,t}l^S_{j,t}\sqrt{2}\tau}\right)^2}+\right.\right.\right.$$
$$\left.\left.\left.\frac{\left(R^l_{t+1}l^S_{j,t}-R^{dS}_{t+1}l^S_{j,t}+R^{dS}_{t+1}e^S_{j,t}\right)}{2}\left[1+\text{erf}\left(\frac{\left(R^l_{t+1}-R^{dS}_{t+1}\right)Q_t}{\sigma^S_{j,t}\sqrt{2}\tau}+\frac{R^{dS}_{t+1}e^S_{j,t}Q_t}{\sigma^S_{j,t}l^S_{j,t}\sqrt{2}\tau}\right)\right]\right)\right]\right\}.$$

The value function of the shadow bank can be written as:

$$V^S_{j,t}=\Gamma^S_{j,t}+\theta E_t\left\{\Lambda_{t,t+1}V^S_{j,t+1}\right\}=E_t\sum_{i=0}^{\infty}\theta^i\Lambda_{t,t+i}\Gamma^S_{j,t+i}.$$

In Appendix C I show that $\Gamma^S_{j,t}$ is increasing in σ_t. Accordingly, each term in the summation above is increasing in σ_t. Then, after applying the law of iterative expectations, it is easy to see that $V^S_{j,t}$ is also increasing in σ_t.

D.3 Aggregating Shadow Banks

$$V^S_{j,t}=\max_{l^S_{j,t+i}}\hat{E}_t\left\{\sum_{i=0}^{\infty}\left(1-\theta\right)\theta^i\Lambda_{t,t+1+i}\omega^S_{j,t+1+i}\right\}$$

subject to

$$\begin{aligned}
\upsilon_{j,t} l_{j,t}^S + \eta_{j,t} e_t^S &\geq \lambda l_{j,t}^S, \\
\omega_{j,t+1+i}^S &= \max\left[\left(R_{t+1+i}^l + \bar{\sigma}^S \frac{\varepsilon_{j,t+1+i}^S}{Q_{t+i}} - R_{t+1+i}^{dS}\right) l_{j,t+i}^S + R_{t+1+i}^{dS} e_{t+i}^S, 0\right].
\end{aligned}$$

Note that the value function depends on one variable only. Thus each banker faces the same problem and chooses the same amount of loans, i.e. $l_{j,t}^S = l_t^S$. The sufficient condition for generating this result is that starting amount of equity is the same across all shadow bankers. Otherwise, depending on the realization of the idiosyncratic shock, each banker will be subject to the financial friction of different strength. In fact, those bankers who receive a relatively large favorable idiosyncratic shock and continue operating in the shadow banking sector have more "skin in the game" and thus can borrow more cheaply. The equity fund that shares equity across all bankers is key to justify my analysis within the representative agent framework.

Moreover, $d_{j,t+1}^S = l_{j,t}^S - e_{j,t}^S = l_t^S - e_t^S = d_{t+1}^S$. It implies that there is a representative shadow banker.

Define:

$$\nu_t = E_t \left\{ (1-\theta)\Lambda_{t,t+1} \left[\bar{\sigma}^S \frac{\tau_S}{Q_t\sqrt{2\pi}} e^{-\left(\frac{\left(R_{t+1}^l - R_{t+1}^{dS}\right) Q_t}{\bar{\sigma}^S \sqrt{2}\tau_S} + \frac{R_{t+1}^{dS} Q_t}{\phi_t^S \bar{\sigma}^S \sqrt{2}\tau_S} \right)^2} + \right.\right.$$
$$\left.\left. \frac{\left(\left(R_{t+1}^l - R_{t+1}^{dS}\right)\right)}{2} \left[1 + \operatorname{erf}\left(\frac{\left(R_{t+1}^l - R_{t+1}^{dS}\right) Q_t}{\bar{\sigma}^S \sqrt{2}\tau_S} + \frac{R_{t+1}^{dS} Q_t}{\phi_t^S \bar{\sigma}^S \sqrt{2}\tau_S} \right)\right]\right] + \theta \Lambda_{t,t+1} \frac{l_{t+1}^S}{l_t^S} \nu_{t+1} \right\}$$

and

$$\eta_t = E_t \left\{ (1-\theta)\Lambda_{t,t+1} \left[\frac{\left(R_{t+1}^{dS}\right)}{2} \left[1 + \operatorname{erf}\left(\frac{\left(R_{t+1}^l - R_{t+1}^{dS}\right) Q_t}{\bar{\sigma}^S \sqrt{2}\tau_S} + \frac{R_{t+1}^{dS} Q_t}{\phi_t^S \bar{\sigma}^S \sqrt{2}\tau_S} \right)\right] + \right.\right.$$
$$\left.\left. \theta \Lambda_{t,t+1} \frac{e_{t+1}^S}{e_t^S} \eta_{t+1} \right]\right\}.$$

Therefore, the incentive constraint can be written as

$$l_t^S = \frac{\eta_t}{\lambda - \upsilon_t} e_t^S = \phi_t^S e_t^S.$$

The terms described in ν_t and η_t evolve according to:

$$z_{t+1} \equiv \frac{e^S_{t+1}}{e^S_t} = \left[\bar{\sigma}^S \phi^S_t \frac{\tau_S}{Q_t\sqrt{2\pi}} e^{-\left(\frac{\left(R^l_{t+1}-R^{dS}_{t+1}\right)Q_t}{\bar{\sigma}^S\sqrt{2}\tau_S}+\frac{R^{dS}_{t+1}Q_t}{\phi^S_t\bar{\sigma}^S\sqrt{2}\tau_S}\right)^2} + \right.$$

$$\left.\frac{\left(R^l_{t+1}\phi^S_t - R^{dS}_{t+1}\phi^S_t + R^{dS}_{t+1}\right)}{2}\left[1+\operatorname{erf}\left(\frac{\left(R^l_{t+1}-R^{dS}_{t+1}\right)Q_t}{\bar{\sigma}^S\sqrt{2}\tau_S}+\frac{R^{dS}_{t+1}Q_t}{\phi^S_t\bar{\sigma}^S\sqrt{2}\tau_S}\right)\right]\right],$$

$$x_{t+1} \equiv \frac{l^S_{t+1}}{l^S_t} = \frac{\phi^S_{t+1}e^S_{t+1}}{\phi^S_t e^S_t}.$$

E The Regulated Bank's Problem

E.1 First-Order Conditions

The objective function can be written recursively

$$V^R_{j,t} = \max_{l^R_{j,t},e^R_{j,t},\sigma^R_{j,t}} \left\{z_{j,t} + \hat{E}_t\left\{\Lambda_{t,t+1}V^R_{j,t+1}\right\}\right\},$$

subject to

$$\begin{aligned}
e^R_{j,t} &\geq \gamma^R_t l^R_{j,t}, \\
z_{j,t} &= \max\left[\left(R^l_t + \sigma^R_{j,t-1}\frac{\varepsilon^R_{j,t}}{Q_{t-1}} - R^{dR}_{t-1}\right)l^R_{j,t-1} + R^{dR}_{t-1}e^R_{j,t-1}, 0\right] - e_{j,t}, \\
l^R_{j,t} &\geq 0, \\
\underline{\sigma}^R \leq \sigma^R_{j,t} &\leq \bar{\sigma}^R.
\end{aligned}$$

Define

$$J\left(S_{t-1}\right) = \max_{l^R_{j,t},d^R_{j,t},e^R_{j,t},\sigma^R_{j,t}} \left\{-e_{j,t} + \hat{E}_t\left[\Lambda_{t,t+1}V^R_{j,t+1}\left(l_{j,t},e_{j,t},\sigma_{j,t},S_t\right)\right]\right\},$$

where S_t is the aggregate state of the economy. Then

$$V^R_{j,t}\left(l_{j,t-1},e_{j,t-1},\sigma_{j,t-1},S_{t-1}\right) = \max\left[\left(R^l_t + \sigma^R_{j,t-1}\frac{\varepsilon^R_{j,t}}{Q_{t-1}} - R^{dR}_{t-1}\right)l^R_{j,t-1} + R^{dR}_{t-1}e^R_{j,t-1}, 0\right] + J\left(S_t\right).$$

Since the first term in the summation on the right hand side is taken as given and $\{\varepsilon_{j,t}\}$

is i.i.d. sequence of random variables, each banker faces the same maximization problem

$$J\left(S_{t-1}\right)=\max_{l_t,e_t,\sigma_t}\left\{-e_t+E_t\left[\Lambda_{t,t+1}\left(\int_{\varepsilon^*_{t+1}}^{\infty}\left(\left(R^l_{t+1}+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}-R^d_t\right)l_t+R^d_te_t\right)\,\mathrm{d}G\left(\varepsilon_{t+1}\right)+J\left(S_t\right)\right)\right]\right\},$$

$$\begin{aligned}&e_t\geq\gamma_tl_t,\\&\left(R^l_{t+1}+\sigma_t\frac{\varepsilon^*_{t+1}}{Q_t}-R^d_t\right)l_t+R^d_te_t=0,\\&l_t\geq 0,\\&\underline{\sigma}\leq\sigma_t\leq\bar{\sigma}.\end{aligned}$$

Append the Lagrange multiplier χ_{1t} to the constraint $e_t\geq\gamma l_t$ and χ_{2t} to the constraint $l_t\geq 0$. Conditional on the optimal choice of σ_t, the first-order conditions are:

$$E_t\left[\Lambda_{t,t+1}\overbrace{\left(\left(R^l_{t+1}+\sigma_t\frac{\varepsilon^*_{t+1}}{Q_t}\right)l_t-R^d_t\left(l_t-e_t\right)\right)}^{=0}\cdot\frac{\partial}{\partial l_t}\left(\frac{R^d_t\left(l_t-e_t\right)Q_t}{\sigma_tl_t}-\frac{R^l_{t+1}Q_t}{\sigma_t}\right)\right]-\gamma\chi_{1t}+\chi_{2t}+$$

$$E_t\left[\int_{\frac{R^d_t(l_t-e_t)Q_t}{\sigma_tl_t}-\frac{R^l_{t+1}Q_t}{\sigma_t}}^{\infty}\Lambda_{t,t+1}\frac{\partial}{\partial l_t}\left(\left(R^l_{t+1}+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}\right)l_t-R^d_t\left(l_t-e_t\right)\right)\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}\right]=0,$$

$$E_t\left[\Lambda_{t,t+1}\overbrace{\left(\left(R^l_{t+1}+\sigma_t\frac{\varepsilon^*_{t+1}}{Q_t}\right)l_t-R^d_t\left(l_t-e_t\right)\right)}^{=0}\cdot\frac{\partial}{\partial e_t}\left(\frac{R^d_t\left(l_t-e_t\right)Q_t}{\sigma_tl_t}-\frac{R^l_{t+1}Q_t}{\sigma_t}\right)\right]-1+\chi_{1t}+$$

$$E_t\left[\int_{\frac{R^d_t(l_t-e_t)Q_t}{\sigma_tl_t}-\frac{R^l_{t+1}Q_t}{\sigma_t}}^{\infty}\Lambda_{t,t+1}\frac{\partial}{\partial e_t}\left(\left(R^l_{t+1}+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}\right)l_t-R^d_t\left(l_t-e_t\right)\right)\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}\right]=0,$$

respectively. I use the Leibniz integral rule above to find the partial derivatives. Note that the first term is zero in the differentiation because the upper limit of the integral does not depend on any of the choice variables. Complementary slackness:

$$\chi_{1t}\left(e_t - \gamma_t l_t\right) = 0,$$
$$\chi_{2t} l_t = 0,$$
$$e_t - \gamma_t l_t \geq 0,$$
$$l_t \geq 0,$$
$$\chi_{1t} \geq 0,$$
$$\chi_{2t} \geq 0,$$

Envelope theorem:

$$J_l'\left(l_{t-1}, e_{t-1}, \sigma_{t-1}\right) = 0,$$
$$J_e'\left(l_{t-1}, e_{t-1}, \sigma_{t-1}\right) = 0.$$

Using the expressions of the integrals from Appendix A, the FOCs can be described by

$$\chi_{2t} + E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\frac{\sigma_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{R_t^d\left(1-\frac{e_t}{l_t}\right)Q_t - R_{t+1}^l Q_t + \xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)^2} + \right.\right.$$
$$\left.\left. + \left(\frac{R_{t+1}^l - \sigma_t\xi - R_t^d}{2}\right)\left[1 - \operatorname{erf}\left(\frac{R_t^d\left(1-\frac{e_t}{l_t}\right)Q_t - R_{t+1}^l Q_t + \xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right]\right]\right\} = \gamma\chi_{1t},$$
$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[R_t^d\frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{R_t^d\left(l_t - e_t\right)Q_t - R_{t+1}^l l_t Q_t + \xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right)\right]\right\} - 1 + \chi_{1t} = 0.$$

E.2 Risk Choice

Denoting Γ_t^R as expected dividends:

$$\Gamma_t^R = -e_t + E_t\left\{\Lambda_{t,t+1}\left[\left(\frac{\sigma_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{R_t^d\left(1-\frac{e_t}{l_t}\right)Q_t - R_{t+1}^l Q_t + \xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)^2} + \right.\right.\right.$$
$$\left.\left.\left.\frac{\left(R_{t+1}^l l_t - R_t^d l_t + R_t^d e_t\right)}{2}\left[1 + \operatorname{erf}\left(\frac{R_t^d\left(1-\frac{e_t}{l_t}\right)Q_t - R_{t+1}^l Q_t + \xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right]\right)\right]\right\}.$$

The objective function of the regulated bank can be written as:

$$J\left(S_{t-1}\right)=\Gamma_t^R+E_t\left\{\Lambda_{t,t+1}J\left(S_t\right)\right\}=E_t\sum_{i=0}^{\infty}\Lambda_{t,t+i}\Gamma_{j,t+i}^S.$$

In Appendix C I show that Γ_t^R is convex in σ_t. Accordingly, each term in the summation above is convex in σ_t. Then, after applying the law of iterative expectations, it is easy to see that $J\left(S_{t-1}\right)$ is also convex in σ_t.

This result guarantees that all the intermediate values $\underline{\sigma}<\sigma_t<\bar{\sigma}$, which may result from the first-order conditions with respect to σ_t, are not optimal.

E.3 Proof of Proposition 1

Equations (41) and (42) can be expressed as

$$\beta E_t\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{i,t+1}^e=1-\frac{\mu_{i,t}}{\lambda_{ct}},$$

where $i\in\{s,r\}$ denotes the type of equity. Using this expression, substitute for 1 in the bank's FOC with respect to e_t. Therefore,

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[R_t^d\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{R_t^d\left(l_t-e_t\right)-R_{t+1}^l l_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2\tau}}\right)\right)\right]-R_{i,t+1}^e\right\}-\frac{\mu_{i,t}}{\lambda_{ct}}+\chi_{1t}=0.$$

Since the range of the erf function is between -1 and 1, i.e.$-1\leq\operatorname{erf}(x)\leq 1$, I know that the following expression is between Ψ_1^* and Ψ_2^*:

$$\Psi_1^*\leq E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[R_t^d\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{R_t^d\left(l_t-e_t\right)-R_{t+1}^l l_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2\tau}}\right)\right)-R_{t+1}^e\right]\right\}\leq\Psi_2^*,$$

where

$$\Psi_1^*=E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[0-R_{i,t+1}^e\right]\right\},$$
$$\Psi_2^*=E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[R_t^d-R_{i,t+1}^e\right]\right\}.$$

Summing up equation (38) with $E_t\beta\lambda_{ct+1}R_{i,t+1}^e+\mu_{i,t}=\lambda_{ct}$ that comes from the household's FOCs with respect to $e_{i,t}$ for each $i\in\{s,r\}$, one gets:

$$E_t\left\{\beta\lambda_{ct+1}\left[R_t^d-R_{i,t+1}^e\right]\right\}=-\sigma_0 d_t^{-\sigma_d}+\mu_{i,t}.$$

The Lagrange multiplier on the households budget constraint, λ_{ct}, is positive. It reflects the fact that the budget constraint always binds given the standard assumptions on the preferences (Inada conditions). Furthermore, $\sigma_0 d_t^{-\sigma_d} > 0$ under the usual (and mild) assumptions on the preferences for liquidity. Dividing the latest expression by λ_{ct} it can be transformed into:

$$\underbrace{E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[R_t^d - R_{i,t+1}^e \right] \right\}}_{=\Psi_2^*} - \frac{\mu_{i,t}}{\lambda_{ct}} = -\frac{\sigma_0 d_t^{-\sigma_d}}{\lambda_{ct}} < 0.$$

Thus, $\Psi_2^* < \frac{\mu_{i,t}}{\lambda_{ct}}$. Therefore,

$$E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[R_t^d \frac{1}{2} \left(1 - \operatorname{erf} \left(\frac{R_t^d (l_t - e_t) - R_{t+1}^l l_t + \xi \sigma_t l_t}{\sigma_t l_t \sqrt{2\tau}} \right) \right) \right] - R_{i,t+1}^e \right\} - \frac{\mu_{i,t}}{\lambda_{ct}} + \chi_{1t} =$$
$$0 < \Psi_2^* - \frac{\mu_{i,t}}{\lambda_{ct}} + \chi_{1t} < \frac{\mu_{i,t}}{\lambda_{ct}} - \frac{\mu_{i,t}}{\lambda_{ct}} + \chi_{1t} = \chi_{1t}.$$

Hence, $\chi_{1t} > 0$. □

E.4 Aggregation of Regulated Banks

The expected dividend that enters the value function of regulated bank j,

$$\begin{aligned} &\mathit{\Omega} \left(\mu_t, \sigma_{j,t};\ l_{j,t} \right) = E_t \left\{ \Lambda_{t,t+1} z_{j,t+1} \right\} = \\ &- E_t \left\{ \Lambda_{t,t+1} \left(1 - \gamma_t \right) l_{j,t}^R \right\} + E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} l_{j,t}^R \left[\sigma_{j,t}^R \frac{\tau_R}{Q_t \sqrt{2\pi}} e^{-\left(\frac{R_t^{dR}(1-\gamma_t)Q_t - R_{t+1}^l Q_t + \xi_R \sigma_{j,t}^R}{\sigma_{j,t}^R \sqrt{2}\tau_R} \right)^2} + \right. \right. \\ &\left. \left. \frac{1}{2} \left(R_{t+1}^l - \sigma_{j,t}^R \xi_R - (1 - \gamma_t) R_t^{dR} \right) \left[1 - \operatorname{erf} \left(\frac{R_t^{dR} (1 - \gamma_t) Q_t - R_{t+1}^l Q_t + \xi_R \sigma_{j,t}^R}{\sigma_{j,t}^R \sqrt{2}\tau_R} \right) \right] \right] \right\}, \end{aligned} \tag{34}$$

is linear in loans. This implies that the value function is also linear in loans.

F The Firm's Problem

Within each class, firms are financed by two types of banks indexed by $i = R, S$. Without loss of generality, I can index firms such that firm j financed by bank i, with $j \in [0, \nu_t^i]$, is risky; firm j financed by bank i, with $j \in [\nu_t^i, 1]$, is safe. Firms financed by regulated banks and firms financed by shadow banks face the same maximization problems, so I omit the index of the type of bank in the next equations.

F.1 Safe firms

For $i \in [\nu_t, 1]$:

The optimality condition on the choice of labor by the safe firm is:

$$H_{i,t+1} = (1-\alpha)\,\frac{Y_{i,t+1}}{W_{t+1}} = (1-\alpha)\,\frac{A_{t+1}K_{i,t+1}^{\alpha}H_{i,t+1}^{1-\alpha}}{W_{t+1}}. \tag{35}$$

Accordingly, the safe firm's Lagrangian is:

$$\begin{aligned}\mathcal{L}^{\text{safe}} =& E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[A_{t+1}K_{i,t+1}^{\alpha}H_{i,t+1}^{1-\alpha} + (1-\delta)Q_{t+1}K_{i,t+1} - W_{t+1}H_{i,t+1} - R_{t+1}^{l}l_{i,t}\right]\right\} \\ &+ \lambda_{Ht}^{s}E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[(1-\alpha)\,\frac{A_{t+1}K_{i,t+1}^{\alpha}H_{i,t+1}^{1-\alpha}}{W_{t+1}} - H_{i,t+1}\right]\right\} + \lambda_{lt}^{s}\left(l_{i,t} - Q_tK_{i,t+1}\right).\end{aligned}$$

Notice that there is no expectation operator on the Lagrange multipliers because those constraints hold under every state of nature. The problem implies the following first-order conditions

$$\begin{aligned}\frac{\partial\mathcal{L}^{\text{safe}}}{\partial l_{i,t}} &= -E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{i,t+1}^{l}\right\} + \lambda_{lt}^{s} = 0, \\ \frac{\partial\mathcal{L}^{\text{safe}}}{\partial K_{i,t+1}} &= E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\alpha\frac{Y_{i,t+1}}{K_{i,t+1}} + (1-\delta)Q_{t+1}\right]\right\} + \\ &\quad \lambda_{Ht}^{s}\left(1-\alpha\right)\alpha E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\frac{Y_{i,t+1}}{W_{t+1}K_{i,t+1}}\right\} - \lambda_{lt}^{s}Q_t = 0, \\ \frac{\partial\mathcal{L}^{\text{safe}}}{\partial H_{i,t+1}} &= (1-\alpha)\,\frac{A_{t+1}K_{i,t+1}^{\alpha}H_{i,t+1}^{1-\alpha}}{W_{t+1}} - W_{t+1} + \lambda_{Ht}^{s}\left[(1-\alpha)^2\,\frac{Y_{i,t+1}}{H_{i,t+1}W_{t+1}} - 1\right] = 0.\end{aligned}$$

Combining $\frac{\partial\mathcal{L}^{\text{safe}}}{\partial H_{i,t+1}} = 0$ with equation (35) yields $\lambda_{Ht}^{s} = 0$. Plugging $\frac{\partial\mathcal{L}^{\text{safe}}}{\partial l_{i,t}} = 0$ into $\frac{\partial\mathcal{L}^{\text{safe}}}{\partial K_{i,t+1}}$ for λ_{lt}^{s}, I get

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{i,t+1}^{l}\right\}Q_t = E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\alpha\frac{Y_{i,t+1}}{K_{i,t+1}} + (1-\delta)Q_{t+1}\right]\right\}.$$

Consider the zero-profit condition of the safe firm under all states of nature. Due to equation (35) I have:

$$A_{t+1}K_{i,t+1}^{\alpha}H_{i,t+1}^{1-\alpha} = \alpha Y_{i,t+1} + W_{t+1}H_{i,t+1}.$$

Substituting this result for $Y_{i,t+1}$ into $\pi_{i,t+1} = 0$ and using $Q_tK_{i,t+1} = l_{i,t}$ yield:

$$\alpha E_tA_{t+1}\left(\frac{K_{i,t+1}}{H_{i,t+1}}\right)^{\alpha-1}K_{i,t+1} + (1-\delta)Q_{t+1}K_{i,t+1} - R_{i,t+1}^{l}Q_tK_{i,t+1} = 0.$$

Thus, $R^l_{i,t+1}Q_t = \alpha E_t A_{t+1}\left(\frac{K_{i,t+1}}{H_{i,t+1}}\right)^{\alpha-1} + (1-\delta)Q_{t+1}$ under all states of nature. This condition implies

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^l_{i,t+1}\right\}Q_t = E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\alpha\frac{Y_{i,t+1}}{K_{i,t+1}} + (1-\delta)Q_{t+1}\right]\right\}.$$

F.2 Risky Firms

For $j \in [0, \nu_t]$:

The optimality condition on the choice of labor by the risky firm is:

$$H_{j,t+1} = \left(\frac{(1-\alpha)A_{t+1}}{W_{t+1}}\right)^{1/\alpha} K_{j,t+1}. \tag{36}$$

Accordingly, the risky firm's Lagrangian is:

$$\begin{aligned}\mathcal{L}^{\text{risky}} =& E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[A_{t+1}K^{\alpha}_{j,t+1}H^{1-\alpha}_{j,t+1} + \varepsilon_{j,t+1}K_{j,t+1} + (1-\delta)Q_{t+1}K_{j,t+1} - W_{t+1}H_{j,t+1} - R^l_{j,t+1}l_{j,t}\right]\right\} + \\ & \lambda^r_{Ht}E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\left(\frac{(1-\alpha)A_{t+1}}{W_{t+1}}\right)^{1/\alpha}K_{j,t+1} - H_{j,t+1}\right]\right\} + \lambda^r_{lt}\left(l_{j,t} - Q_tK_{j,t+1}\right).\end{aligned}$$

Notice that there is no expectation operator on the Lagrange multipliers because those constraints hold under every state of nature. The problem implies the following first-order conditions

$$\begin{aligned}\frac{\partial\mathcal{L}^{\text{risky}}}{\partial l_{j,t}} &= -E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^l_{j,t+1}\right\} + \lambda^r_{lt} = 0,\\ \frac{\partial\mathcal{L}^{\text{risky}}}{\partial K_{j,t+1}} &= E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\alpha A_{t+1}\left(\frac{K_{j,t+1}}{H_{j,t+1}}\right)^{\alpha-1} + \varepsilon_{j,t+1} + (1-\delta)Q_{t+1}\right]\right\} + \\ &\quad \lambda^r_{Ht}E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\frac{(1-\alpha)A_{t+1}}{W_{t+1}}\right)^{1/\alpha}\right\} - \lambda^r_{lt}Q_t = 0,\\ \frac{\partial\mathcal{L}^{\text{risky}}}{\partial H_{j,t+1}} &= (1-\alpha)A_{t+1}\left(\frac{K_{j,t+1}}{H_{j,t+1}}\right)^{\alpha} - W_{t+1} + \lambda^r_{Ht}[-1] = 0.\end{aligned}$$

Combining $\frac{\partial\mathcal{L}^{\text{risky}}}{\partial H_{j,t+1}} = 0$ with equation (36) yields $\lambda^r_{Ht} = 0$. Plugging $\frac{\partial\mathcal{L}^{\text{risky}}}{\partial l_{j,t}} = 0$ into $\frac{\partial\mathcal{L}^{\text{risky}}}{\partial K_{j,t+1}}$ for λ^r_{lt}, I get

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^l_{j,t+1}\right\}Q_t = E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\alpha A_{t+1}\left(\frac{K_{j,t+1}}{H_{j,t+1}}\right)^{\alpha-1} + (1-\delta)Q_{t+1} + \varepsilon_{j,t+1}\right]\right\}.$$

Combining equation (35) with equation (36) results in:

$$\frac{K_{i,t+1}}{H_{i,t+1}} = \frac{K_{j,t+1}}{H_{j,t+1}} = \frac{K_{t+1}}{H_{t+1}} \tag{37}$$

under all states of nature. But remember that

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^l_{i,t+1}\right\}Q_t = E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\alpha A_{t+1}\left(\frac{K_{i,t+1}}{H_{i,t+1}}\right)^{\alpha-1} + (1-\delta)Q_{t+1}\right]\right\} =$$
$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\alpha A_{t+1}\left(\frac{K_{t+1}}{H_{t+1}}\right)^{\alpha-1} + (1-\delta)Q_{t+1}\right]\right\} = E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^l_{t+1}\right\}Q_t.$$

Therefore

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^l_{j,t+1}\right\}Q_t = E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[R^l_{t+1}Q_t + \varepsilon_{j,t+1}\right]\right\}.$$

Consider the zero-profit condition of the risky firm under all states of nature. Due to equation (36) I have:

$$A_{t+1}K^{\alpha}_{j,t+1}H^{1-\alpha}_{j,t+1} = \alpha A_{t+1}\left(\frac{K_{j,t+1}}{H_{j,t+1}}\right)^{\alpha-1}K_{j,t+1} + W_{t+1}H_{j,t+1}.$$

Substituting this result for $Y_{j,t+1}$ into $\pi_{j,t+1} = 0$, using $Q_tK_{j,t+1} = l_{j,t}$ and equation (37) yield:

$$R^l_{j,t+1}Q_tK_{j,t+1} + \varepsilon_{j,t+1}K_{j,t+1} - R^l_{j,t+1}Q_tK_{j,t+1} = 0.$$

Thus, $R^l_{j,t+1}Q_t = R^l_{t+1}Q_t + \varepsilon_{j,t+1}$ under all states of nature. This condition implies

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^l_{j,t+1}\right\}Q_t = E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[R^l_{t+1} + \varepsilon_{j,t+1}\right]\right\}.$$

Therefore, at the time the bank j is making a loan to the safe firm, the bank expects to receive the total returns on safe loans equal $E_tR^l_{t+1}l^s_{j,t}$ and the total returns on risky loans equal $E_tR^l_{j,t+1}l^r_{j,t} = E_t\left(R^l_{t+1} + \frac{\varepsilon_{j,t+1}}{Q_t}\right)l^r_{j,t}$. Summing them up yields

$$E_tR^l_{t+1}l^s_{j,t} + E_tR^l_{j,t+1}l^r_{j,t} =$$
$$E_tR^l_{t+1}\left(1-\sigma_{j,t}\right)l_{j,t} + E_t\left(R^l_{t+1} + \frac{\varepsilon_{j,t+1}}{Q_t}\right)\sigma_{j,t}l_{j,t} = \left(R^l_{t+1} + \sigma_{j,t}\frac{\varepsilon_{j,t+1}}{Q_t}\right)l_{j,t},$$

where $l_{j,t} = l^s_{j,t} + l^r_{j,t}$ and $l^s_{j,t} = \left(1-\sigma_{j,t}\right)l_{j,t}$. So, the returns in the bank's maximization problem are consistent with the firm's problem.

F.3 Aggregating Across Firms

Here I show that I can aggregate individual firms into two representative firms. Let denote $K_{i,t}^{j}$ the capital chosen by firm i that is financed by borrowing from bank j. In this notation, the equation (37) is written as

$$\frac{K_{i,t+1}^{j}}{H_{i,t+1}^{j}} = \frac{K_{t+1}}{H_{t+1}},$$

for all $j \in [0,1]$ and $i \in [0,1]$.

Define the following objects: Let $K_{S,t+1}^{S} = \int_{\mu_t}^{1} \int_{\nu_t}^{1} K_{i,t+1}^{j} didj$ be the total capital allocated to the safe technology and financed by borrowing from the banks that choose a fraction $\underline{\sigma}$ of risky projects. Let $K_{S,t+1}^{R} = \int_{0}^{\mu_t} \int_{\nu_t}^{1} K_{i,t+1}^{j} didj$ be the total capital allocated to the safe technology and financed by borrowing from the banks that choose a fraction $\bar{\sigma}$ of risky projects. Thus,

$$K_{S,t+1} = K_{S,t+1}^{S} + K_{S,t+1}^{R},$$

where $K_{S,t+1}$ is the total capital allocated to the safe technology.

Let $K_{R,t+1}^{S} = \int_{\mu_t}^{1} \int_{0}^{\nu_t} K_{i,t+1}^{j} didj$ be the total capital allocated to the risky technology and financed by borrowing from the banks that choose a fraction $\underline{\sigma}$ of risky projects. Let $K_{R,t+1}^{R} = \int_{0}^{\mu_t} \int_{0}^{\nu_t} K_{i,t+1}^{j} didj$ be the total capital allocated to the safe technology and financed by borrowing from the banks that choose a fraction $\bar{\sigma}$ of risky projects. Thus,

$$K_{R,t+1} = K_{R,t+1}^{S} + K_{R,t+1}^{R},$$

where $K_{R,t+1}$ is the total capital allocated to the risky technology. The same upper and lower case notation applies to labor, i. e. $H_{S,t+1}^{S} = \int_{\mu_t}^{1} \int_{\nu_t}^{1} H_{i,t+1}^{j} didj$; $H_{S,t+1}^{R} = \int_{0}^{\mu_t} \int_{\nu_t}^{1} H_{i,t+1}^{j} didj$; $H_{R,t+1}^{S} = \int_{\mu_t}^{1} \int_{0}^{\nu_t} H_{i,t+1}^{j} didj$; $H_{R,t+1}^{R} = \int_{0}^{\mu_t} \int_{0}^{\nu_t} H_{i,t+1}^{j} didj$.

Safe representative firm:

$$Y_t^S = \int\limits_0^1 \int\limits_{\nu_{t-1}}^1 F\left(K_{i,t}^j, H_{i,t}^j\right) didj = \int\limits_0^1 \int\limits_{\nu_{t-1}}^1 A_t \left(K_{i,t}^j\right)^\alpha \left(H_{i,t}^j\right)^{1-\alpha} didj =$$

$$\int\limits_0^1 \int\limits_{\nu_{t-1}}^1 A_t \left[F_{K_{i,t}^j}\left(K_{i,t}^j, H_{i,t}^j\right) K_{i,t}^j + F_{H_{i,t}^j}\left(K_{i,t}^j, H_{i,t}^j\right) H_{i,t}^j\right] didj =$$

$$\int\limits_0^1 \int\limits_{\nu_{t-1}}^1 A_t \left[f_{K_{i,t}^j}\left(\frac{K_{i,t}^j}{H_{i,t}^j}\right) K_{i,t}^j + f_{H_{i,t}^j}\left(\frac{K_{i,t}^j}{H_{i,t}^j}\right) H_{i,t}^j\right] didj =$$

$$\int\limits_0^1 \int\limits_{\nu_{t-1}}^1 A_t \left[f_{K_t}\left(\frac{K_t}{H_t}\right) K_{i,t}^j + f_{H_t}\left(\frac{K_t}{H_t}\right) H_{i,t}^j\right] didj =$$

$$\int\limits_0^{\mu_{t-1}} \int\limits_{\nu_{t-1}}^1 A_t \left[f_{K_t}\left(\frac{K_t}{H_t}\right) K_{i,t}^j + f_{H_t}\left(\frac{K_t}{H_t}\right) H_{i,t}^j\right] didj +$$

$$\int\limits_{\mu_{t-1}}^1 \int\limits_{\nu_{t-1}}^1 A_t \left[f_{K_t}\left(\frac{K_t}{H_t}\right) K_{i,t}^j + f_{H_t}\left(\frac{K_t}{H_t}\right) H_{i,t}^j\right] didj =$$

$$A_t \left[f_{K_t}\left(\frac{K_t}{H_t}\right) K_{S,t}^R + f_{H_t}\left(\frac{K_t}{H_t}\right) H_{S,t}^R\right] + A_t \left[f_{K_t}\left(\frac{K_t}{H_t}\right) K_{S,t}^S + f_{H_t}\left(\frac{K_t}{H_t}\right) H_{S,t}^S\right] =$$

$$A_t \left[f_{K_t}\left(\frac{K_t}{H_t}\right) \left(K_{S,t}^R + K_{S,t}^S\right) + f_{H_t}\left(\frac{K_t}{H_t}\right) \left(H_{S,t}^R + H_{S,t}^S\right)\right] =$$

$$A_t \left[f_{K_t}\left(\frac{K_t}{H_t}\right) K_{S,t} + f_{H_t}\left(\frac{K_t}{H_t}\right) H_{S,t}\right] = A_t \left(K_{S,t}\right)^\alpha \left(H_{S,t}\right)^{1-\alpha},$$

where $K_{S,t} = \int_{\mu_{t-1}}^1 \left(1 - \underline{\sigma}\right) l_{j,t-1} dj + \int_0^{\mu_{t-1}} \left(1 - \bar{\sigma}\right) l_{j,t-1} dj$.

Risky representative firm:

$$Y_t^R = \int\limits_0^1 \int\limits_0^{\nu_{t-1}} \left[F\left(K_{i,t}^j, H_{i,t}^j\right) + \varepsilon_{i,t} K_{i,t}^j\right] didj = A_t \left(K_{R,t}\right)^\alpha \left(H_{R,t}\right)^{1-\alpha} +$$

$$\int\limits_0^1 \int\limits_0^{\nu_{t-1}} \varepsilon_{i,t} K_{i,t}^j didj = A_t \left(K_{R,t}\right)^\alpha \left(H_{R,t}\right)^{1-\alpha} + \left(-\xi\right) K_{R,t} = A_t \left(K_{R,t}\right)^\alpha \left(H_{R,t}\right)^{1-\alpha} - \xi K_{R,t},$$

where $K_{R,t} = \int_{\mu_{t-1}}^1 \underline{\sigma} l_{j,t-1} dj + \int_0^{\mu_{t-1}} \bar{\sigma} l_{j,t-1} dj$.

G The Government

The government levies the tax to fully compensate for the loss to the deposit insurance fund due to rescue of defaulted banks.

$$T_t = -\int_{-\infty}^{\varepsilon_t^*} \left(\left(R_t^l + \frac{\sigma_{t-1}\varepsilon_t}{Q_{t-1}} \right) l_{t-1} - R_{t-1}^d d_{t-1} \right) \mathrm{d}G(\varepsilon_t) =$$

$$\frac{\sigma_{t-1} l_{t-1}}{Q_{t-1}} \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{R_{t-1}^d (1-\gamma_{t-1}) Q_{t-1} - R_t^l Q_{t-1} + \xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau} \right)^2} -$$

$$\frac{1}{2} \left(R_t^l l_{t-1} - \frac{\sigma_{t-1}\xi}{Q_{t-1}} l_{t-1} - R_{t-1}^d d_{t-1} \right) \left[1 + \operatorname{erf} \left(\frac{R_{t-1}^d (1-\gamma_{t-1}) Q_{t-1} - R_t^l Q_{t-1} + \xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau} \right) \right].$$

H Household

To express $x_t^\star$, substitute $\xi = 0$ into the expression derived in Appendix B. Thus,

$$x_t^\star = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{\left(R_t^l - R_t^{dS} \right) Q_{t-1}}{\bar{\sigma}^S \sqrt{2}\tau_S} + \frac{R_t^{dS} Q_{t-1}}{\bar{\sigma}^S \phi_{t-1}^S \sqrt{2}\tau_S} \right) \right],$$

When a shadow bank defaults, it liquidates its assets to partially reimburse depositors. In such situation, the deposit fund gets

$$\Pi_t^S = (1 - x_t^\star) R_t^{dS} d_t^S + \int_{-\infty}^{\varepsilon_t^*} \left(\left(R_t^l + \bar{\sigma}^S \varepsilon_t - R_t^{dS} \right) l_{t-1}^S + R_t^{dS} e_{t-1}^S \right) \mathrm{d}G(\varepsilon_t).$$

It comprises the earnings that the fund would get if the defaulted deposits would pay in full plus the loss it makes due to non-repayment of the full deposit rate. Rewriting it as follows:

$$\Pi_t^S = (1 - x_t^\star) R_t^{dS} d_t^S + \left(\left(R_t^l l_{t-1}^S - R_t^{dS} d_t^S \right) - \int_{\varepsilon_t^*}^{\infty} \left(\left(R_t^l + \bar{\sigma}^S \varepsilon_t - R_t^{dS} \right) l_{t-1}^S + R_t^{dS} e_{t-1}^S \right) \mathrm{d}G(\varepsilon_t) \right) =$$

$$(1 - x_t^\star) R_t^{dS} d_t^S + R_t^l l_{t-1}^S - R_t^{dS} d_t^S - l_{t-1}^S \bar{\sigma}^S \frac{\tau_S}{Q_{t-1}\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^l - R_t^{dS} \right) Q_{t-1}}{\bar{\sigma}^S \sqrt{2}\tau_S} + \frac{R_t^{dS} Q_{t-1}}{\phi_{t-1}^S \bar{\sigma}^S \sqrt{2}\tau_S} \right)^2} -$$

$$\left(R_t^l l_{t-1}^S - R_t^{dS} d_t^S \right) x_t^\star = R_t^l l_{t-1}^S (1 - x_t^\star) - l_{t-1}^S \bar{\sigma}^S \frac{\tau_S}{Q_{t-1}\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^l - R_t^{dS} \right) Q_{t-1}}{\bar{\sigma}^S \sqrt{2}\tau_S} + \frac{R_t^{dS} Q_{t-1}}{\phi_{t-1}^S \bar{\sigma}^S \sqrt{2}\tau_S} \right)^2}.$$

The household's first-order conditions are:

$$
\begin{aligned}
D^R: &\quad \sigma_0 \Psi'_{D_t^R}\left(D_t^R, D_t^S\right) - \lambda_{ct} + \beta E_t \lambda_{ct+1} R_t^{dR} = 0, && (38)\\
D^S: &\quad \sigma_0 \Psi'_{D_t^S}\left(D_t^R, D_t^S\right) - \lambda_{ct} + \beta E_t \left\{\lambda_{ct+1} x_{t+1}^{\star}\right\} R_{t+1}^{dS} = 0, && (39)\\
R: &\quad C_t^{-\sigma_c} - \lambda_{ct} = 0, && (40)\\
E_s^R: &\quad -\lambda_{ct} + \beta E_t \left\{\lambda_{ct+1} R_{s,t+1}^{eR}\right\} + \mu_{s,t}^R = 0, && (41)\\
E_r^R: &\quad -\lambda_{ct} + \beta E_t \left\{\lambda_{ct+1} R_{r,t+1}^{eR}\right\} + \mu_{r,t}^R = 0, && (42)\\
E_H^S: &\quad -\lambda_{ct} + \beta E_t \left\{\lambda_{ct+1} R_{t+1}^{eS}\right\} = 0, && (43)
\end{aligned}
$$

where λ_{ct}, $\mu_{s,t}^R$ and $\mu_{r,t}^R$ are the Lagrangian multipliers for the budget constraint and the two non-negativity equity constraints.

Equations (38) and (39) are the Euler equations for the consumption-deposit choice. The term $\frac{\partial \Psi\left(D_t^R, D_t^S\right)}{\partial D}$ can be interpreted as a liquidity premium for holding each type of deposits. Equation (40) expresses the marginal utility of consumption. Equations (41), (42) and (43) determine the state-contingent required rates of return on three types of equity.

I Data and Calibration

I.1 Measuring the Sizes of the Banking Sectors

The measures of the banking sectors come from the Z.1 release of the Financial Accounts of the United States (Flow of Funds) as of 2022:Q1. The nominal size of regulated banks is estimated with FOF code 'FL704190005.Q'. The nominal measure of shadow banks sums the following FOF codes: 'FL792150005.Q', 'FL634090005.Q', 'FL893169175.Q', 'FL663070675.Q', 'FL674122005.Q' and 'FL614190005.Q'. To measure them in real terms, I divide my estimates by GDP implicit price deflator from FRED ('USAGDPDEFQISMEI').

I.2 Calibration

Accounting for the banking costs, the expressions of loan returns become:

$$
\begin{aligned}
\omega_{t+1}^S &= \max\left[\left(R_{t+1}^l + \sigma_t^S \frac{\varepsilon_{t+1}^S}{Q_t} - f^S\right) l_t^S - R_{t+1}^{dS} d_{t+1}^S, 0\right],\\
e_{t+1}^{j,R} &= \max\left[\left(R_{t+1}^l + \sigma_t^{j,R} \frac{\varepsilon_{t+1}^R}{Q_t} - f^R\right) l_t^{j,R} + R_t^{dR} d_t^{j,R}, 0\right] - z_{t+1}^j,
\end{aligned} \tag{44}
$$

where $j = s, r$ stands for the safe and risky representative regulated bank, respectively.

To calibrate the variance of the idiosyncratic shock τ_R^2, I link the production function of the risky firm to the production function of the safe firm that has a preexisting debt. Remember that the next period returns to safe and risky loans are given by

$$\begin{aligned} R^l_{t+1} &= \frac{\alpha A_{t+1}}{Q_t}\left(\frac{K_{t+1}}{H_{t+1}}\right)^{\alpha-1} + (1-\delta)\frac{Q_{t+1}}{Q_t}, \\ R^{lr}_{t+1} &= R^l_{t+1} + \sigma_{\mathrm{RF}}\frac{\varepsilon_{t+1}}{Q_t}, \end{aligned}$$

respectively. The parameter σ_{RF} captures the exposure to the idiosyncratic shock. The risky bank that finances the maximum share of risky projects earns

$$\Omega^{risky}_{t+1} = R^{lr}_{t+1} Q_t K^r_{t+1}.$$

It comprises EBITDA and what the bank makes or loses by selling the capital to capital producers. The safe bank with preexisting debt earns

$$\Omega^{safe}_{t+1} = R^l_{t+1} Q_t \left(K_{t+1} + B_t\right) - Q_t B_t R^B_t = \left(R^l_{t+1}\left(1 + \frac{B_t}{K_{t+1}}\right) - \frac{B_t}{K_{t+1}} R^B_t\right) Q_t K_{t+1},$$

where B_t is a predetermined debt, measured in units of capital, and R^B_t is a predetermined interest rate. I equate the conditional variances of the returns to loans

$$Var_t\left(R^{lr}_{t+1}\right) = Var_t\left(R^l_{t+1}\left(1 + \frac{B_t}{K_{t+1}}\right) - \frac{B_t}{K_{t+1}} R^B_t\right)$$

to find the variance of the idiosyncratic shock that matches $\frac{\text{Debt}}{\text{EBITDA}} = 7$. Note that

$$\begin{aligned} Var_t\left(R^{lr}_{t+1}\right) &= Var_t\left(R^l_{t+1}\right) + \left(\frac{\sigma_{\mathrm{RF}}}{Q_t}\right)^2 \tau_R^2, \\ Var_t\left(R^l_{t+1}\left(1 + \frac{B_t}{K_{t+1}}\right) - \frac{B_t}{K_{t+1}} R^B_t\right) &= \left(1 + \frac{B_t}{K_{t+1}}\right)^2 Var_t\left(R^l_{t+1}\right), \end{aligned}$$

where K_{t+1} is the steady-state level of capital of the safe firms that are financed by regulated banks and $Q_t = 1$.

The conditional variance of the returns on safe loans is given by

$$\begin{aligned} Var_t\left(R^l_{t+1}\right) = \alpha^2\left(\frac{K_{t+1}}{H_{t+1}}\right)^{2\alpha-2} Var_t\left(A_{t+1}\right) + (1-\delta)^2 Var_t\left(Q_{t+1}\right) + \\ 2\alpha\left(\frac{K_{t+1}}{H_{t+1}}\right)^{\alpha-1}(1-\delta) Cov_t\left(A_{t+1}, Q_{t+1}\right). \end{aligned}$$

There is a way to calculate the conditional variance of Q_{t+1} by picking up its process from the optimization problem of capital producers. However, my approach is meant to be suggestive, and I consider that the conditional variance of Q_{t+1} is the same as the conditional variance of the aggregate shock. The covariance term is expected to be positive but I drop it in my calculations because the terms that multiply it are small. The model's counterpart for EBITDA is a total output net of compensation for labor. Thus

$$\frac{\text{Debt}}{\text{EBITDA}} = \frac{B_t}{Y_t^{safe} - W_t H_t^{safe}} = \frac{B_t}{\alpha Y_t^{safe}}.$$

The data analog of σ_{RF} is the share of leveraged loans held by regulated banks. I calibrate σ_{RF} to 45% that comes from the Shared National Credit Report issued by the Fed, OCC, and FDIC.

J Optimal Policies

J.1 Optimal Dynamic Capital Requirements

The Ramsey problem can be written in general terms as follows:

$$W_0 = \max_{\{\tilde{x}_t, \gamma_t\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t U\left(\tilde{x}_{t-1}, \tilde{x}_t, \varsigma_t\right),$$

subject to

$$E_t g\left(x_{t-1}, x_t, x_{t+1}, \varsigma_t\right) = 0,$$

where x_t is a vector of endogenous variables and ς_t is a vector of exogenous variables. The $N \times 1$ vector x_t is partitioned as $x_t = (\tilde{x}_t', \gamma_t)'$. Given the sequence of the policy instrument $\{\gamma_t\}_{t=0}^{\infty}$, the remaining $N-1$ endogenous variables need to satisfy the $N-1$ structural equilibrium conditions which are captured by the vector g.

I propose $\{\gamma_t^*\}_{t=0}^{\infty}$ that is set at the lowest level that prevents regulated banks from financing risky firms in every period following the shock. To understand how I choose it, consider an alternative path $\left\{\gamma_t^A\right\}_{t=0}^{\infty}$ in which $\gamma_t^A = \gamma_t^*$ for $t \neq t_k$ and $\gamma_t^A = \gamma_t^* + \Delta$ for $t = t_k$ where $\Delta \neq 0$. When $\Delta > 0$, $\left\{\gamma_t^A\right\}_{t=0}^{\infty}$ is welfare dominated by $\{\gamma_t^*\}_{t=0}^{\infty}$ because a higher capital requirement in period t_k leads to welfare losses from the reduced amount of liquidity services without altering risk-taking incentives. This holds for any t_k (and any combination of t_k) and does not depend on the size of $\Delta > 0$. When $\Delta < 0$, regulated banks switch into financing socially inefficient risky projects in period t_k under $\left\{\gamma_t^A\right\}_{t=0}^{\infty}$. On the one hand, a decrease in the capital requirement involves a social loss of ξ_R for making loans. On the other hand, it increases liquidity services that enter into the utility function of households directly. The trade-off between these two considerations determines the impact on welfare. For a relatively small decrease in capital requirements (i.e. Δ is close to zero), the former consideration is more important, and thus total welfare falls. In fact, there is an inefficiency loss of approximately $\xi_R Q_t l_{t_k}^{R*}$ from a switch into financing risky loans (where $l_{t_k}^{R*}$ is the amount of loans in period t_k under the path $\{\gamma_t^*\}_{t=0}^{\infty}$), while the effects on liquidity are negligible or even negative due to the inefficiency that encompasses all loans after the switch. For a larger decrease in capital requirements (i.e. for $\Delta < \Delta^*$), the latter consideration starts to dominate. In fact, since households value each unit of liquidity services relatively more than the economy, on average, loses from making risky loans (i.e. $\varsigma_0 > \xi_R$ in the calibration), welfare is decreasing in Δ, and thus the benefits from liquidity services can outweigh the costs of switching into financing risky projects. The gains from liquidity services are maximized when $\gamma_{t_k}^A = 0$. Therefore, I need to compare conditional welfare under $\{\gamma_t^*\}_{t=0}^{\infty}$ against the alternatives that let the capital requirement fall to zero in some periods.

I use a similar numerical procedure developed in Canzoneri et al. (2020) to verify whether corner conditions implying zero capital requirements that maximize the benefits from liquidity provision are dominated by the policy described above. I find that the calculated capital requirements are, in fact, welfare superior to that alternative for each of the shocks considered in the text.

J.2 Adaptation to Occbin

To adapt my framework to OccBin, I break it into three different models. The constraints that capture the signs of the Lagrange multipliers on the loan and equity constraints of regulated banks control the switching from one model to another. Accordingly,

1. The Safe model describes the reference regime with $\mu_t = 0$ in which $\chi_{2,t}^s = 0$, $\chi_{2,t}^r > 0$, $\mu_{s,t}^R = 0$, and $\mu_{r,t}^R > 0$.

2. The Risky model describes the alternative regime with $\mu_t = 1$ in which $\chi_{2,t}^s > 0$, $\chi_{2,t}^r = 0$, $\mu_{s,t}^R > 0$, and $\mu_{r,t}^R = 0$.

3. The Mixed model describes the alternative regime with $0 < \mu_t < 1$ in which $\chi_{2,t}^s > 0$, $\chi_{2,t}^r > 0$, $\mu_{s,t}^R > 0$, and $\mu_{r,t}^R > 0$.

K Capital Quality Shock

K.1 Firms – including capital quality shock

The safe production technology is given by

$$Y^s_{j,t} = A_t \left(\iota_t K_{j,t}\right)^\alpha H_{j,t}^{1-\alpha},$$

for $j \in [\nu_t, 1]$ where A_t is a total factor productivity and ι_t is the quality of capital (so that $\iota_t K_{j,t}$ is the effective amount of capital). The risky production technology is given by

$$Y^r_{j,t} = A_t \left(\iota_t K_{j,t}\right)^\alpha H_{j,t}^{1-\alpha} + \varepsilon_{j,t}\iota_t K_{j,t},$$

for $j \in [0, \nu_t]$ and where $\varepsilon_{j,t}$ is the same idiosyncratic shock described in the main text.

The maximization problem can be expressed as

$$\max_{l_t,\, K_{j,t+1}} E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \max_{H_{j,t+1}} \left(Y_{j,t+1} + (1-\delta) Q_{t+1}\iota_{t+1} K_{j,t+1} - W_{t+1} H_{j,t+1} - R^l_{j,t+1} l_{j,t} \right) \right\}, \quad (45)$$

subject to constraints of the production technology and financing $Q_t K_{j,t+1} = l_{j,t}$. The solution implies that:

$$W_{t+1} \quad = \quad (1-\alpha)\, A_{t+1} \left(\frac{\iota_{t+1} K_{j,t+1}}{H_{j,t+1}} \right)^\alpha . \quad (46)$$

under all states of nature.

The Lagrangian of the risky firm is:

$$\begin{aligned} \mathcal{L}^{\text{risky}} =& E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[A_{t+1} \left(\iota_{t+1} K_{j,t+1}\right)^\alpha H_{j,t+1}^{1-\alpha} + \varepsilon_{j,t+1}\iota_{t+1} K_{j,t+1} + (1-\delta) Q_{t+1}\iota_{t+1} K_{j,t+1} - W_{t+1} H_{j,t+1} - R^l_{j,t+1} l_{j,t} \right] \right\} + \\ & \lambda^r_{Ht} E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[\left(\frac{(1-\alpha)\, A_{t+1}}{W_{t+1}} \right)^{1/\alpha} \iota_{t+1} K_{j,t+1} - H_{j,t+1} \right] \right\} + \lambda^r_{lt} \left(l_{j,t} - Q_t K_{j,t+1} \right). \end{aligned}$$

Notice that there is no expectation operator on the Lagrangian multipliers because those constraints hold under every state of nature. The problem implies the following first-order conditions

$$\begin{aligned}
\frac{\partial \mathcal{L}^{\text{risky}}}{\partial l_{j,t}} &= -E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^l_{j,t+1}\right\}+\lambda^r_{lt}=0,\\
\frac{\partial \mathcal{L}^{\text{risky}}}{\partial K_{j,t+1}} &= E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\alpha A_{t+1}\iota^{\alpha}_{t+1}\left(\frac{K_{j,t+1}}{H_{j,t+1}}\right)^{\alpha-1}+\iota_{t+1}\varepsilon_{j,t+1}+(1-\delta)\iota_{t+1}Q_{t+1}\right]\right\}+\\
&\quad \lambda^r_{Ht}E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\frac{(1-\alpha)A_{t+1}}{W_{t+1}}\right)^{1/\alpha}\iota_{t+1}\right\}-\lambda^r_{lt}Q_t=0,\\
\frac{\partial \mathcal{L}^{\text{risky}}}{\partial H_{j,t+1}} &= (1-\alpha)A_{t+1}\left(\frac{\iota_{t+1}K_{j,t+1}}{H_{j,t+1}}\right)^{\alpha}-W_{t+1}+\lambda^r_{Ht}[-1]=0.
\end{aligned}$$

Combining $\frac{\partial \mathcal{L}^{\text{risky}}}{\partial H_{j,t+1}}=0$ with equation (46) yields $\lambda^r_{Ht}=0$. Then, plugging $\frac{\partial \mathcal{L}^{\text{risky}}}{\partial l_{j,t}}=0$ into $\frac{\partial \mathcal{L}^{\text{risky}}}{\partial K_{j,t+1}}$ for λ^r_{lt}, we get

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^l_{j,t+1}\right\}Q_t=E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\alpha A_{t+1}\iota^{\alpha}_{t+1}\left(\frac{K_{j,t+1}}{H_{j,t+1}}\right)^{\alpha-1}+(1-\delta)\iota_{t+1}Q_{t+1}+\iota_{t+1}\varepsilon_{j,t+1}\right]\right\}.$$

Applying equation (46) for both types of firms results in:

$$\frac{K_{i,t+1}}{H_{i,t+1}}=\frac{K_{j,t+1}}{H_{j,t+1}}=\frac{K_{t+1}}{H_{t+1}} \tag{47}$$

under all states of nature. But remember that

$$\begin{aligned}
&E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^l_{i,t+1}\right\}Q_t=E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\alpha A_{t+1}\iota^{\alpha}_{t+1}\left(\frac{K_{i,t+1}}{H_{i,t+1}}\right)^{\alpha-1}+(1-\delta)\iota_{t+1}Q_{t+1}\right]\right\}=\\
&\qquad E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\alpha A_{t+1}\iota^{\alpha}_{t+1}\left(\frac{K_{t+1}}{H_{t+1}}\right)^{\alpha-1}+(1-\delta)\iota_{t+1}Q_{t+1}\right]\right\}=E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^l_{t+1}\right\}Q_t.
\end{aligned}$$

Therefore

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^l_{j,t+1}\right\}Q_t=E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[R^l_{t+1}Q_t+\iota_{t+1}\varepsilon_{j,t+1}\right]\right\}.$$

The solution implies

$$\begin{aligned}
R^l_t\equiv R^l_{i,t} &= \frac{\alpha A_t\iota^{\alpha}_t}{Q_{t-1}}\left(\frac{K_t}{H_t}\right)^{\alpha-1}+(1-\delta)\iota_t\frac{Q_t}{Q_{t-1}}=\frac{\alpha Y^s_{i,t}}{K_{i,t}Q_{t-1}}+(1-\delta)\iota_t\frac{Q_t}{Q_{t-1}},\\
R^l_{j,t} &= R^l_t+\frac{\iota_t\varepsilon_{j,t}}{Q_{t-1}},
\end{aligned}$$

for all $j\in[0,\nu_t]$ and $i\in[\nu_t,1]$.

K.2 Banks – Including Capital Quality Shock

In case we include capital adjustment costs together with a capital quality shock, the bank's problem is described as follows:

$$\max_{l_t,e_t,\sigma_t} E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\int\limits_{\left(\frac{R_t^d(l_t-e_t)}{\sigma_t l_t}-\frac{R_{t+1}^l}{\sigma_t}\right)\frac{Q_t}{\iota_{t+1}}}^{\infty}\left(\left(R_{t+1}^l+\frac{\sigma_t\iota_{t+1}\varepsilon_{t+1}}{Q_t}\right)l_t-R_t^d\left(l_t-e_t\right)\right)\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}-(1+\kappa)\,e_t\right].$$

Expression of the integral:

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}l_t\left[\frac{\sigma_t\iota_{t+1}}{Q_t}\int\limits_{\left(\frac{R_t^d(l_t-e_t)}{\sigma_t l_t}-\frac{R_{t+1}^l}{\sigma_t}\right)\frac{Q_t}{\iota_{t+1}}}^{\infty}\varepsilon_{t+1}\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}+\right.\right.$$

$$\left.\left.\left(R_{t+1}^l-R_t^d\left(1-\gamma_t\right)\right)\int\limits_{\left(\frac{R_t^d(l_t-e_t)}{\sigma_t l_t}-\frac{R_{t+1}^l}{\sigma_t}\right)\frac{Q_t}{\iota_{t+1}}}^{\infty}\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}\right]\right\}.$$

Break the calculation of the integral into two parts.

$$\int\limits_{\left(\frac{R_t^d(l_t-e_t)}{\sigma_t l_t}-\frac{R_{t+1}^l}{\sigma_t}\right)\frac{Q_t}{\iota_{t+1}}}^{\infty}\varepsilon_{t+1}\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}=$$

Introduce a change in variables to recast the integral in terms of the Standard Normal distribution. Use $v=\frac{\varepsilon_{t+1}+\xi}{\sqrt{2}\tau}$, or equivalently $\varepsilon_{t+1}=v\sqrt{2}\tau-\xi$, and remember that for the change $x=\varphi(t)$, the integral $\int_{\varphi(a)}^{\varphi(b)}f(x)dx$ becomes $\int_a^b f(\varphi(t))\varphi'(t)dt$. Here we use that $dv=\frac{d\varepsilon_{t+1}}{\sqrt{2}\tau}$, so we need to multiply dv by $\sqrt{2}\tau$ to express $d\varepsilon_{t+1}$ in terms of dv. Moreover, we need to transform the lower limit using v. So we need to add ξ to the lower limit of the integral and divide the result by $\sqrt{2}\tau$.

$$\int\limits_{\frac{\left(R_t^d(1-\gamma_t)-R_{t+1}^l\right)Q_t+\sigma_t\iota_{t+1}\xi}{\sigma_t\iota_{t+1}\sqrt{2}\tau}}^{\infty}\left(v\sqrt{2}\tau-\xi\right)\frac{\sqrt{2}\tau}{\sqrt{2\pi\tau^2}}e^{-v^2}\,\mathrm{d}v=\int\limits_{\frac{\left(R_t^d(1-\gamma_t)-R_{t+1}^l\right)Q_t+\sigma_t\iota_{t+1}\xi}{\sigma_t\iota_{t+1}\sqrt{2}\tau}}^{\infty}\left(v\sqrt{2}\tau-\xi\right)\frac{1}{\sqrt{\pi}}e^{-v^2}\,\mathrm{d}v=$$

$$\frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^d(1-\gamma_t)-R_{t+1}^l\right)Q_t+\sigma_t\iota_{t+1}\xi}{\sigma_t\iota_{t+1}\sqrt{2}\tau}\right)^2} - \frac{\xi}{2}\left[1-\text{erf}\left(\frac{\left(R_t^d\left(1-\gamma_t\right)-R_{t+1}^l\right)Q_t+\sigma_t\iota_{t+1}\xi}{\sigma_t\iota_{t+1}\sqrt{2}\tau}\right)\right].$$

Where we used that $\text{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-v^2}$.

Let's express $\int_{\left(\frac{R_t^d(l_t-e_t)}{\sigma_t l_t}-\frac{R_{t+1}^l}{\sigma_t}\right)\frac{Q_t}{\iota_{t+1}}}^{\infty} \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\, \text{d}\varepsilon_{t+1}$ in terms of the error function. Again, use the transformation $v = \frac{\varepsilon_{t+1}+\xi}{\sqrt{2}\tau}$ or $\varepsilon_{t+1} = v\sqrt{2}\tau - \xi$

$$\int\limits_{\left(\frac{R_t^d(l_t-e_t)}{\sigma_t l_t}-\frac{R_{t+1}^l}{\sigma_t}\right)\frac{Q_t}{\iota_{t+1}}}^{\infty} \frac{\sqrt{2}\tau}{\sqrt{2\pi\tau^2}} e^{-v^2}\,\text{d}v = \frac{1}{\sqrt{\pi}} \int\limits_{\left(\frac{R_t^d(l_t-e_t)}{\sigma_t l_t}-\frac{R_{t+1}^l}{\sigma_t}\right)\frac{Q_t}{\iota_{t+1}}}^{\infty} e^{-v^2}\,\text{d}v =$$
$$\frac{1}{2}\left(1-\text{erf}\left(\left(\frac{R_t^d\left(l_t-e_t\right)}{\sigma_t l_t}-\frac{R_{t+1}^l}{\sigma_t}\right)\frac{Q_t}{\iota_{t+1}}\right)\right).$$

Therefore,

$$\Omega(\mu_t,\sigma_t;\, l_t,\, d_t,\, e_t) = E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}} l_t\left[\frac{\left(R_{t+1}^l - R_t^d\left(1-\gamma_t\right)\right)}{2}\left[1-\text{erf}\left(\frac{\left(R_t^d\left(1-\gamma_t\right)-R_{t+1}^l\right)Q_t+\sigma_t\iota_{t+1}\xi}{\sigma_t\iota_{t+1}\sqrt{2}\tau}\right)\right]+\right.\right.$$
$$\left.\left.\frac{\sigma_t\iota_{t+1}}{Q_t}\left[\frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^d(1-\gamma_t)-R_{t+1}^l\right)Q_t+\sigma_t\iota_{t+1}\xi}{\sigma_t\iota_{t+1}\sqrt{2}\tau}\right)^2} - \frac{\xi}{2}\left[1-\text{erf}\left(\frac{\left(R_t^d\left(1-\gamma_t\right)-R_{t+1}^l\right)Q_t+\sigma_t\iota_{t+1}\xi}{\sigma_t\iota_{t+1}\sqrt{2}\tau}\right)\right]\right]\right]\right\}.$$

L Derivation of Δ

$$E\sum_{t=0}^{\infty}\beta^t\left[\frac{(1-\Delta)C_{\text{opt},t}^{1-\sigma_c}-1}{1-\sigma_c}+\sigma_0\Psi\left(D_{\text{opt},t}^R,D_{\text{opt},t}^S\right)\right]=$$
$$E\sum_{t=0}^{\infty}\beta^t\left[\frac{C_{\text{rule},t}^{1-\sigma_c}-1}{1-\sigma_c}+\sigma_0\Psi\left(D_{\text{rule}_j,t}^R,D_{\text{rule}_j,t}^S\right)\right],$$
$$E\sum_{t=0}^{\infty}\beta^t\left[\frac{C_t^{1-\varrho_c}-1}{1-\varrho_c}+\varsigma_0\frac{D_t^{1-\varsigma_d}-1}{1-\varsigma_d}\right]$$

Let $Welf^{opt}$ be the welfare level attained under the optimal policy and $Welf^{rule}$ be the welfare level attained under a particular rule. For given paths of consumption and deposits, we are interested in sizing a permanent tax Δ applied to the consumption utility stream under the optimal rule such that the level of welfare under the optimal policy with the tax is equal to the level of welfare under the suboptimal rule. Thus,

$$E\sum_{t=0}^{\infty}\beta^t\left[\frac{((1-\Delta)C_{\text{opt},t})^{1-\sigma_c}-1}{1-\sigma_c}+\sigma_0\Psi\left(D^R_{\text{opt},t},D^S_{\text{opt},t}\right)\right]=Welf^{rule},$$

which can be rewritten as

$$E\sum_{t=0}^{\infty}\beta^t\left[\frac{(1-\Delta)^{1-\sigma_c}\left(C_{\text{opt},t}\right)^{1-\sigma_c}-1}{1-\sigma_c}+\sigma_0\Psi\left(D^R_{\text{opt},t},D^S_{\text{opt},t}\right)\right]=Welf^{rule},$$

Taking out $(1-\Delta)^{1-\sigma_c}$:

$$E\sum_{t=0}^{\infty}\beta^t\left[(1-\Delta)^{1-\sigma_c}\frac{\left(C_{\text{opt},t}\right)^{1-\sigma_c}-1}{1-\sigma_c}+\sigma_0\Psi\left(D^R_{\text{opt},t},D^S_{\text{opt},t}\right)+\frac{(1-\Delta)^{1-\sigma_c}-1}{1-\sigma_c}\right]=Welf^{rule},$$

It can be further re-written as follows:

$$E\sum_{t=0}^{\infty}\beta^t\left[\left((1-\Delta)^{1-\sigma_c}-1\right)\frac{\left(C_{\text{opt},t}\right)^{1-\sigma_c}-1}{1-\sigma_c}+\frac{\left(C_{\text{opt},t}\right)^{1-\sigma_c}-1}{1-\sigma_c}+\right.$$
$$\left.\sigma_0\Psi\left(D^R_{\text{opt},t},D^S_{\text{opt},t}\right)+\frac{(1-\Delta)^{1-\sigma_c}-1}{1-\sigma_c}\right]=Welf^{rule},$$

Let denote $Welf_C^{opt}=E\sum\limits_{t=0}^{\infty}\beta^t\frac{\left(C_{\text{opt},t}\right)^{1-\sigma_c}-1}{1-\sigma_c}$ the welfare from the consumption utility stream attained under the optimal policy. Hence,

$$\left((1-\Delta)^{1-\sigma_c}-1\right)Welf_C^{opt}+Welf^{opt}+\frac{(1-\Delta)^{1-\sigma_c}-1}{(1-\sigma_c)(1-\beta)}=Welf^{rule},$$

where it is used that $\sum\limits_{t=0}^{\infty}\beta^t=\frac{1}{1-\beta}$.

Deriving Δ:

$$\Delta=1-\left(1-\frac{Welf^{opt}-Welf^{rule}}{\left(Welf_C^{opt}+\frac{1}{(1-\beta)(1-\sigma_c)}\right)}\right)^{\frac{1}{1-\sigma_c}}.$$

M Additional Results

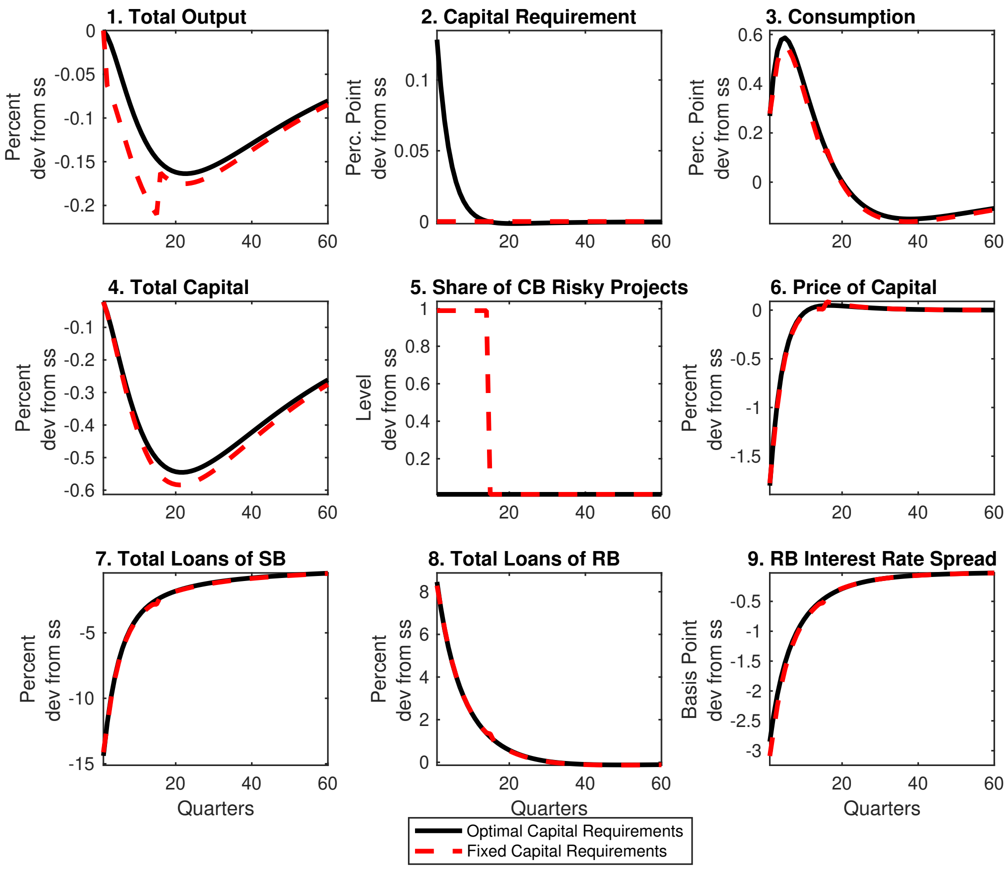

Figure 9. Responses to a Rise in Volatility of Risky Returns in the Shadow Banking Sector: Capital requirements can be still a useful instrument for responding to the purely sectoral shock that occurs in the shadow banking.

Note: This figure plots the responses to a 60 basis point rise in τ_t^S. The shock follows an AR(1) process with an autoregressive coefficient of 0.9. The shock is sized to increase the quarterly default rate of shadow bank loans by 1% on impact. The solid lines show the responses under the optimal capital requirement that is set at the minimum level to prevent excessive risk taking in every period following the shock. The dashed lines show what would happen if the capital requirement were to be held constant at its steady state value. SB and RB stand for shadow and regulated banks, respectively.

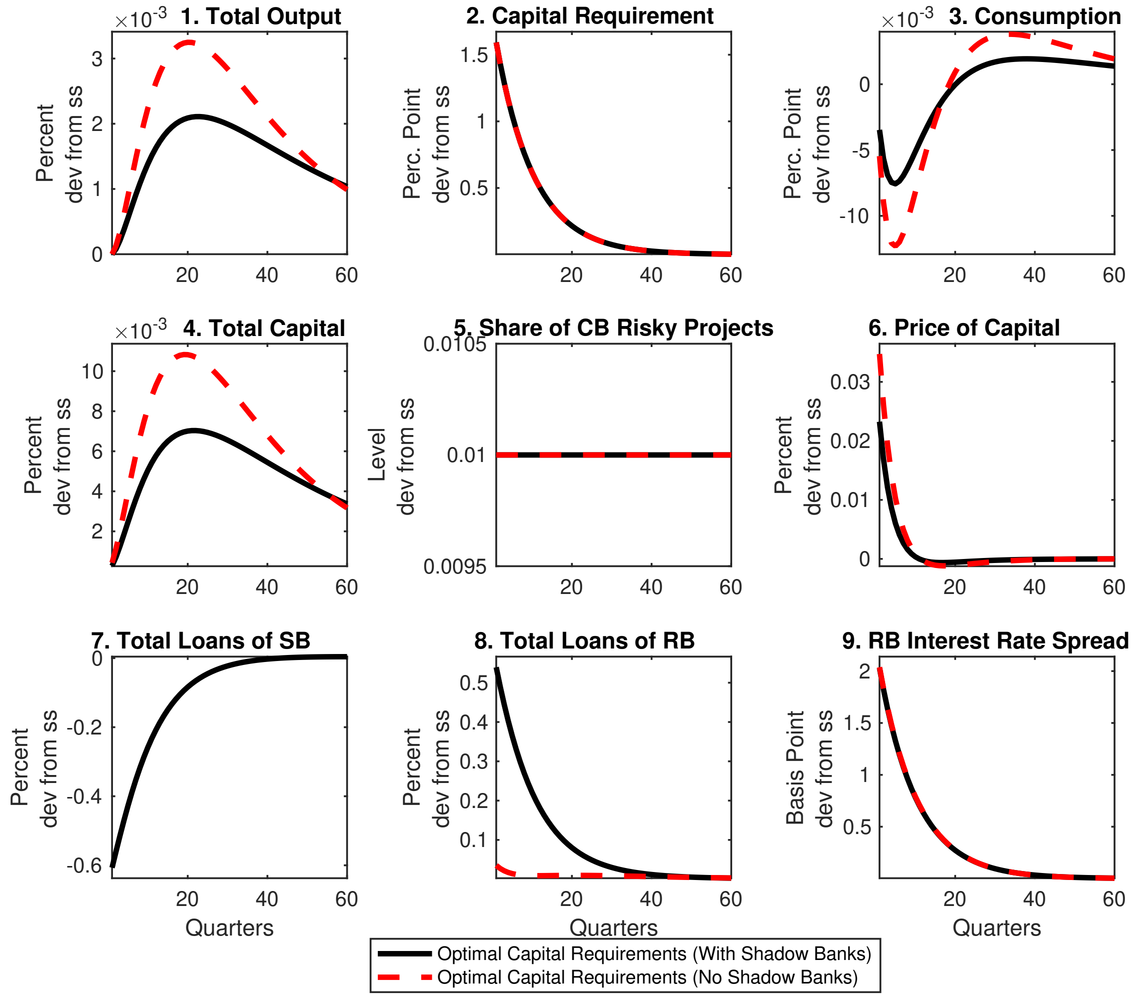

Figure 10. Responses to a Rise in Volatility of Risky Returns in the Regulated Banking Sector: No reintermediation channel.

Note: This figure plots the responses to a 60 basis point rise in τ_t^R under the optimal capital requirement (that is set at the minimum level to prevent excessive risk taking in every period following the shock) for two models. The shock follows an AR(1) process with an autoregressive coefficient of 0.9. The solid lines show the responses of the baseline model with shadow banks. The dashed lines show what would happen in the model without shadow banks (i.e. regulated banks only). SB and RB stand for shadow and regulated banks, respectively.

N Additional Tables and Figures

Table 6: Comparison of Two Types of Banks.

Features / Bank	Deposit Insurance	Risk-Taking
Regulated	Full	Endogenous: $\underline{\sigma}^R \leq \sigma_t \leq \bar{\sigma}^R$
Shadow	No	Endogenous: $\underline{\sigma}^S \leq \sigma_t \leq \bar{\sigma}^S$

Features / Bank	Regulation	Life Span
Regulated	Capital Requirement	Infinitely-lived
Shadow	No	Finite horizon: continues with probability θ

Features / Bank	Modeling Frictions
Regulated	Moral hazard associated with deposit insurance and limited liability
Shadow	Moral hazard associated with limited liability and costly enforcement

Features / Bank	Loan portfolio
Regulated	Returns to Safe projects: $\left(1-\sigma_t^R\right) R_{t+1}^l$ Returns to Risky projects: $\sigma_t^R \left(R_{t+1}^l + \frac{\varepsilon_{t+1}^R}{Q_t}\right)$ where $\varepsilon_{t+1}^R \sim N\left(-\xi_R, \tau_R^2\right)$ for $\xi_R > 0$
Shadow	Returns to Safe projects: $\left(1-\sigma_t^S\right) R_{t+1}^l$ Returns to Risky projects: $\sigma_t^S \left(R_{t+1}^l + \frac{\varepsilon_{t+1}^S}{Q_t}\right)$ where $\varepsilon_{t+1}^S \sim N\left(0, \tau_S^2\right)$

Note: This table shows the main differences in the modeling approach of two types of banks.

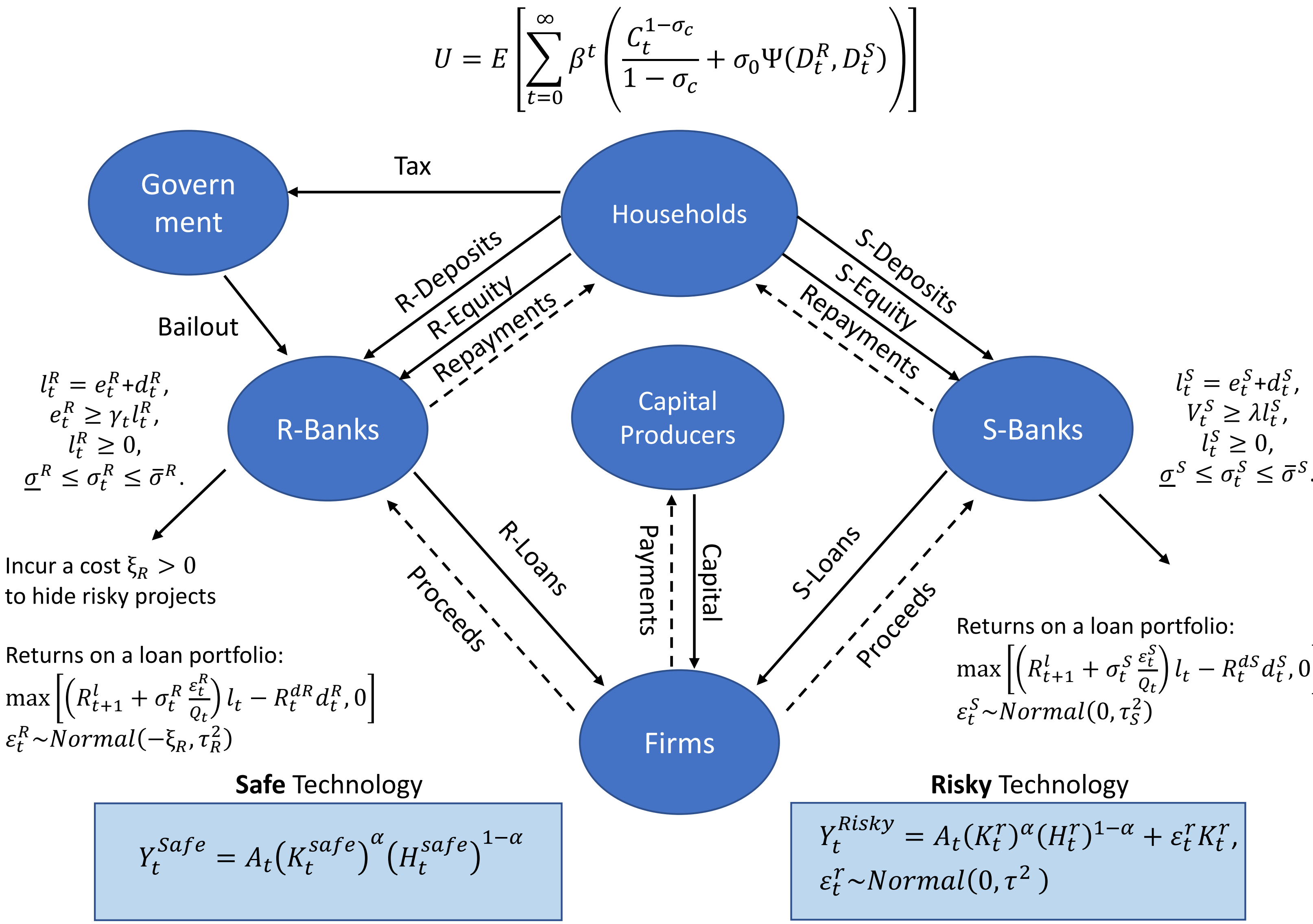

Figure 11. Overview of the Model.

Note: This figure plots the overview of the model.

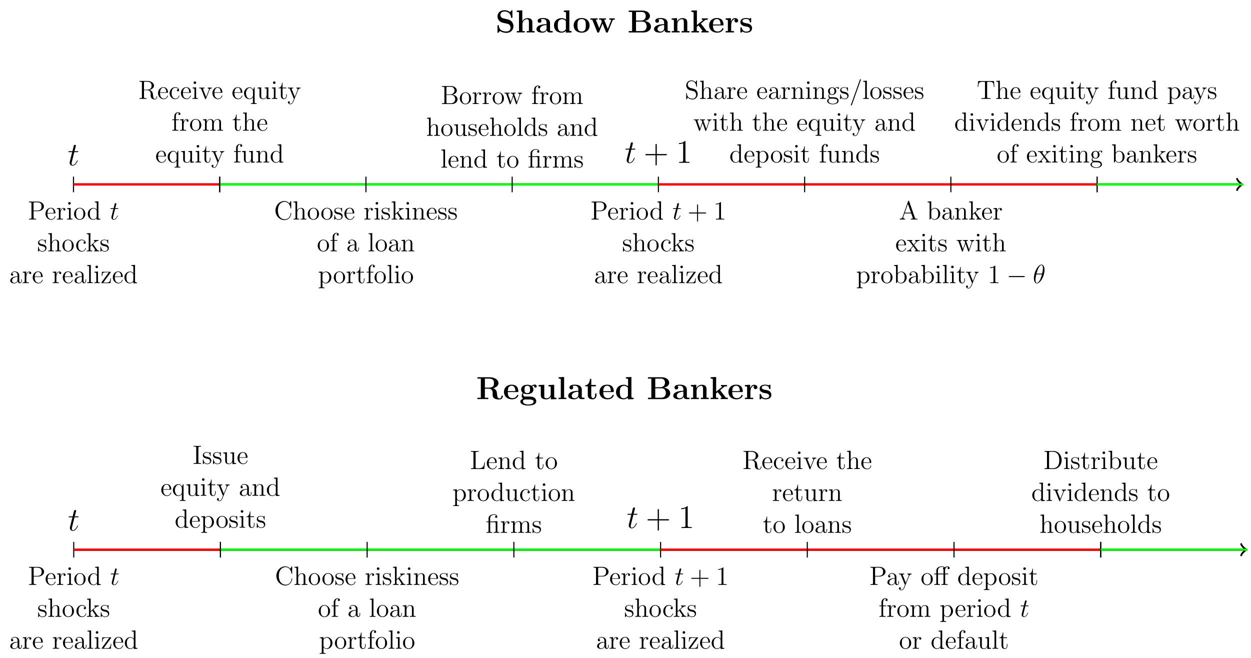

Figure 12. Timeline for Two Types of Bankers.

Note: This figure plots the timelines for two types of banks.