A Static Capital Buffer is Hard To Beat*

Matthew Canzoneri[†] Behzad Diba[‡] Luca Guerrieri[§] Arsenii Mishin[¶]

February 2021[‖]

Abstract

We build a quantitatively relevant macroeconomic model with endogenous risk taking. In our model, deposit insurance and limited liability can lead banks to make risky loans that are socially inefficient. This excessive risk taking can be triggered by aggregate or sectoral shocks that reduce the return on safer loans. Excessive risk taking can be avoided by raising bank capital requirements, but unnecessarily tight requirements lower welfare by reducing liquidity producing bank deposits. Consequently, optimal capital requirements are dynamic (or state contingent). We provide examples in which a Ramsey planner would raise capital requirements: (1) during a downturn caused by a TFP shock; (2) during an expansion caused by an investment specific shock; and (3) during an increase in market volatility that has little effect on the business cycle. In practice, the economy is driven by a constellation of shocks, and the Ramsey policy is beyond the policymaker's ken; so, we also consider implementable policy rules. Basel III guidance calls for increasing capital requirements when the credit to GDP ratio rises, and relaxing them when it falls; this rule does not perform well in our environment. Some rules do mimic the Ramsey policy rather well, but – in the interest of preserving deposits – they may set the reserve requirement too low, triggering a fall into the abyss of a risk-taking episode. Fortunately, slightly elevated static capital requirements (or "buffers") generally do about as well as any implementable rule, without the risk of a Wile E. Coyote moment.

---

*The views expressed here are our own; they do not necessarily reflect the views of the Federal Reserve Board or any other member of its staff. We wish to thank (without implicating) David Arseneau, Skander Van den Heuvel, and Stacey Schreft for helpful comments.

[†]Georgetown University, canzonem@georgetown.edu

[‡]Georgetown University, dibab@georgetown.edu

[§]Federal Reserve Board, luca.guerrieri@frb.gov

[¶]HSE University, aomishin@hse.ru

[‖]An earlier version of this paper was circulated under the title, "Optimal Capital Requirements and Implementable Capital Buffer Rules."

First, do no harm, Hippocrates (5th century BCE)

Mind the cliff, Wile E. Coyote (20th century CE)

1 Introduction

A protracted period of low returns on safe assets followed in the wake of the global financial crisis, and this trend is expected to continue for the foreseeable future. These low returns have raised concerns that financial intermediaries will be tempted to reach for higher yields by taking excessive (or socially inefficient) risks. We formalize these concerns by developing a dynamic macroeconomic model in which limited liability and deposit insurance provide incentives for excessive risk taking: a sudden fall in the returns on safe assets can trigger an extended period of excessive risk taking, with major consequences for consumption, business investment and household welfare. Prudential policy can curb these incentives by raising bank capital requirements; indeed, capital requirements that are high enough will eliminate the incentives entirely. But this will come at the expense of reducing bank deposits, which provide liquidity services to households. An all-knowing Ramsey planner – faced with aggregate and sectoral shocks – would maximize the utility of deposits by lowering capital requirements just short of falling into a risk-taking episode, which would lower household utility by a discrete amount. A less-knowing planner in the real world would try to zip up to the cliff's edge without actually going over. The well-meaning planner runs the risk of a Wile E. Coyote moment, and may be better advised to exercise caution, or do no harm.

In this paper, we will explore this tradeoff both theoretically and quantitatively. More specifically, we calculate optimal policies for dynamic capital requirements, and we study the ability of simple policy rules to mimic them. We provide examples in which a Ramsey Planner would raise capital requirements: (1) during a downturn caused by a TFP shock; (2) during an expansion caused by an investment specific shock; and (3) during an increase in market volatility that has little to do with the business cycle. But in a more realistic setting, where the economy is bombarded by a full constellation of shocks, the Ramsey policy would require too much information to be implementable. So, we also study the ability of simple policy rules to mimic the Ramsey policy. Of particular interest will be the Basel III guidance.

Our DSGE model combines key elements of the literature on financial frictions and macroeconomic stability. Following Van den Heuvel (2008), banks can lend to "safe" firms or "risky" firms. Both kinds of firms are subject to aggregate TFP shocks, but a risky firm is also exposed to an idiosyncratic shock with negative expected value; risky loans are therefore socially inefficient. The only reason a profit maximizing bank would fund a risky firm is that limited liability shields it from downside risk; if the expected return on safe loans is expected to fall, the bank may take a flier on a risky loan. Banks fund their lending by issuing deposits and equity to households. Deposits are the cheaper source of funding since they provide liquidity services, and in addition, government deposit insurance makes them the safe asset. Capital requirements increase funding costs and make banks keep more skin in the game. This effect reduces their temptation to take excessive risks. Van den Heuvel's model does not allow for aggregate economic fluctuations or changes in market volatility. In our model, macroeconomic shocks lead to business cycles, and they can trigger excessive risk taking by decreasing the expected return on safe loans. Market volatility shocks can also trigger excessive risk taking.

As noted above, real world policy makers would probably not be able to implement the full Ramsey policy; the information required would prove daunting. So, we consider simple policy rules that try to mimic the Ramsey policy by responding to just one or two endogenous variables. To this end, we use the simulated method of moments to calibrate our model's dynamic structure, which in turn allows us to calculate Ramsey dynamic capital requirements when the model economy is driven by a full constellation of shocks. We generate model data in that stochastic environment, and we regress the optimal capital requirements on candidate sets of endogenous variables. Some simple rules capture the optimal capital requirements rather well; that is, they have an R-square statistic close to 1, at least for some calibrations. However, most of these rules fall into the risk-taking trap with a frequency that we can calculate. Slightly elevated static capital requirements (or "buffers") avoid the Wile E. Coyote moments, and generally do about as well as any implementable policy rule on deposits.

The Basel III accords advocated a cyclical capital buffer: during credit booms (or increases in the credit-to-GDP ratio), capital requirements would be tightened; during contractions they could be loosened. These prescriptions – which we will call the "Basel rule" – sound sensible, and they should be implementable in practice. But in our model, the Basel rule does not come close to mimicking the Ramsey policy; other simple rules, and small static buffers, can do better.

Literature Review:

A number of contributions to the literature [e.g. Kashyap and Stein (2004), Repullo and Suarez (2013), Gersbach and Rochet (2017), Davydiuk (2017), and Malherbe (2020)] highlight alternative considerations that seem relevant for addressing the pro-cyclical bias of Basel II guidelines and the counter-cyclical buffers of Basel III.[1] The alternative consid-

[1] The procyclical bias of Basel II guidelines is attributed to risk-based capital requirements, which effectively tighten during recessions as the default risk on bank assets increases.

erations include capital scarcity, funding frictions, risk shifting, and pecuniary externalities. These models focus on the effects of capital requirements on the volume of bank credit. By contrast, our model focuses on the composition of bank credit; we have in mind risk-taking decisions like the choice between prime and subprime mortgages before the 2007-2009 financial crisis, or more recently, participation in syndicated loans to highly leveraged firms. We do not intend to take issue with the importance of risks associated with bank leverage (and the volume of bank credit). High leverage can, for example, increase the risk of bank runs in environments like the models of Angeloni and Faia (2013), Gertler and Kiyotaki (2015), Elenev et al. (2018), Faria-e-Castro (2019), and Gertler et al. (2020). Our focus is complementary to the emphasis on leverage and the volume of credit in much of the literature.

Schularick and Taylor (2012), Jordà et al. (2016), and Mian et al. (2017) present empirical evidence that large crises tend to follow rapid credit expansions. The result is often interpreted in favor of the Basel-III style rules. At the same time, there are papers that question the causal interpretation of that empirical evidence. For example, Gomes et al. (2018) develop a model in which bank credit is, by design, independent of output and investment, but nevertheless credit expansions predict output declines. Gomes et al. (2018) share our emphasis on risks that arise from endogenous changes in the composition of bank credit, but not our focus on what this implies for optimal capital requirements. Mendicino et al. (2018) also share our focus on the composition of bank credit. They study the welfare effects of simple rules reacting to default risk in a model with patient and impatient borrowers.

The papers by Martinez-Miera and Suarez (2014), Collard et al. (2017) and Begenau (2019) also examine capital requirements from a perspective similar to ours, but they don't share our focus on cyclical variation in optimal capital requirements. Martinez-Miera and Suarez (2014) develop a model with systemic risk abstracting from aggregate shocks. Collard et al. (2017) focus on interactions of optimal monetary and prudential policies, in a setting that keeps bank failures off the equilibrium path. Begenau (2019) develops a quantitative business-cycle model to determine the optimal level of a constant capital requirement. Van den Heuvel (2019) extends his earlier work (Van den Heuvel, 2008) to quantify the welfare cost of liquidity regulations as well as capital requirements.

The rest of the paper proceeds as follows. Section 2 describes the model. Section 3 discusses the model's calibration, including the choice of steady-state capital requirements. Section 4 describes our numerical methods for the model solutions. Section 5 discusses the Ramsey Policy we take as optimal. Section 6 presents the responses to different shocks and discusses the Ramsey policy for capital requirements. Section 7 considers some simple implementable rules. And Section 8 concludes.

2 The Model

Our model extends a standard RBC model to include banks that enjoy limited liability and government deposit insurance. These are the main features that allow for excessive, or socially inefficient, risk taking, and of course the RBC framework allows for macroeconomic shocks that cause business cycles. Our model consists of households, banks, nonfinancial firms, and a government whose sole purpose is to provide bank deposit insurance. Banks are at the heart of our model, but the exposition is smoother if we begin with the less exciting firms and households.

But first, a note on notation: There are a measure one continua of households, banks and non-financial firms. In what follows, small letters denote individual households, banks or firms; capital letters represent aggregate values. Safe firms (defined below) carry a superscript s; risky firms carry a superscript r.

2.1 Non-Financial Firms

Non-financial firms are competitive and earn zero profits. There are goods producing firms and capital producing firms. We begin with the former.

2.1.1 Goods Producing Firms:

Firms live for just two periods. A firm born in period t, obtains a bank loan, l_t^f, to buy the capital, k_{t+1}, that it will use for production in period $t+1$; so,

$$l_t^f = Q_t k_{t+1}, \tag{1}$$

where Q_t is the price of capital (or the price of investment). The ex-post return on the loan is $R_{t+1} l_t^f = R_{t+1} Q_t k_{t+1}$, where we shall soon see that R_{t+1} is the rate of return on capital ownership. So, these bank loans might be better described as equity positions.

There is a continuum of firms of measure 1. But the firms come in two types: "safe" firms face only aggregate shocks, while "risky" firms face both aggregate shocks and idiosyncratic shocks.

In period $t+1$, a safe firm hires labor, h_{t+1}^s, to produce

$$y_{t+1}^s = A_{t+1} (k_{t+1}^s)^{\alpha} (h_{t+1}^s)^{1-\alpha}, \tag{2}$$

where A_{t+1} is an aggregate TFP shock. When a safe firm takes the loan in period t, it knows that the firm will hire the optimal h_{t+1}^s next period. So, the safe firm chooses $l_t^{f,s}$ and k_{t+1}^sin period t, and then h_{t+1}^s in period $t+1$, to

$$\max_{l_t^{f,s},k_{t+1}^s} E_t \left\{ \max_{h_{t+1}^s} \left[y_{t+1}^s + (1-\delta)Q_{t+1}k_{t+1}^s - W_{t+1}h_{t+1}^s - R_{t+1}^s l_t^{f,s} \right] \right\} \tag{3}$$

where δ is the capital depreciation rate, and W_{t+1} is the real wage rate. This maximization is subject to (1) and (2). The first order conditions for this maximization problem imply

$$E_t R_{t+1}^s = \alpha E_t \left\{ \frac{A_{t+1}}{Q_t} \left(\frac{h_{t+1}^s}{k_{t+1}^s} \right)^{1-\alpha} + (1-\delta)\frac{Q_{t+1}}{Q_t} \right\}, \tag{4}$$

where the first term within the brackets is the rental rate on a unit of capital, and the second term is the capital gain on a non-depreciated unit of capital.

A risky firm employs the technology $y_{t+1}^r = A_{t+1} \left(k_{t+1}^r\right)^\alpha \left(h_{t+1}^r\right)^{1-\alpha} + \varepsilon_{t+1}k_{t+1}^r$, where ε_{t+1} is an idiosyncratic shock that follows a Normal distribution G with a negative mean, $-\xi$, and standard deviation τ:[2]

$$\begin{aligned} &\text{PDF of } \varepsilon_{t+1}, \quad g(\varepsilon_{t+1}) = \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \\ &\text{CDF of } \varepsilon_{t+1}, \quad G(\varepsilon_{t+1}) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\varepsilon_{t+1}+\xi}{\tau\sqrt{2}}\right)\right] \end{aligned} \tag{5}$$

The risky firm chooses $l_t^{f,r}$ and k_{t+1}^r, and then h_{t+1}^r, to

$$\max_{l_t^{f,r},k_{t+1}^r} E_t \left\{ \max_{h_{t+1}^r} \left[y_{t+1}^r + (1-\delta)Q_{t+1}k_{t+1}^r - W_{t+1}h_{t+1}^r - R_{t+1}^r l_t^{f,r} \right] \right\} \tag{6}$$

subject to the analogous constraints. The first order conditions for this maximization, the zero profit condition for firms, and equation (8) below, imply

$$E_t R_{t+1}^r = E_t R_{t+1}^s - \frac{\xi}{Q_t}. \tag{7}$$

So the idiosyncratic shock lowers the expected value, and increases the variance, of the return on a loan to a risky firm. Risky loans are socially inefficient, or in our language, excessively risky.

Note finally that the marginal product of labor for safe and risky firms is $(1-\alpha)A(k_{t+1}^i/h_{t+1}^i)^\alpha$ where i denotes the type of firm ($i \in \{s,r\}$). Labor is mobile across firms, and both types of firms face the same real wage rate. So, the first order conditions for labor in period $t+1$ imply the capital labor ratios equalize across sectors.

[2] $\text{erf}(x) = \frac{1}{\sqrt{\pi}}\int_{-x}^{x} e^{-v^2}dv = \frac{2}{\sqrt{\pi}}\int_0^x e^{-v^2}dv$.

$$k^r_{t+1}/h^r_{t+1} = k^s_{t+1}/h^s_{t+1}. \tag{8}$$

Appendix B.3 provides details on aggregation across firms; there we show that there is a representative safe firm that produces

$$Y^s_{t+1} = A_{t+1}(K^s_{t+1})^{\alpha}(H^s_{t+1})^{1-\alpha}, \tag{9}$$

and also a representative risky firm that produces

$$Y^r_{t+1} = A_{t+1}\left(K^r_{t+1}\right)^{\alpha}\left(H^r_{t+1}\right)^{1-\alpha} - \xi K^r_{t+1}. \tag{10}$$

2.1.2 Capital Producing Firms

At the end of period t, goods producing firms sell their capital to competitive capital producing firms. Letting I^g_t denote gross investment, the evolution of capital follows

$$I_t = \eta_t\left[1 - \frac{\phi}{2}\left(\frac{I^g_t}{I^g_{t-1}} - 1\right)^2\right] I^g_t, \tag{11}$$

where η_t is an investment specific technology shock, and ϕ is a measure of the severity of investment adjustment costs. The aggregate capital stock evolves according to

$$K^s_{t+1} + K^r_{t+1} = I_t + (1-\delta)\left(K^s_t + K^r_t\right). \tag{12}$$

The capital producing firms are owned by households, and solve the problem

$$\max_{I^g_{t+i}} E_t \sum_{i=0}^{\infty} \psi_{t,t+i}\left\{Q_{t+i}\eta_{t+i}\left[1 - \frac{\phi}{2}\left(\frac{I^g_{t+i}}{I^g_{t+i-1}} - 1\right)^2\right] I^g_{t+i} - I^g_{t+i}\right\}, \tag{13}$$

where $\psi_{t,t+i} = \beta\frac{\lambda_{ct+i}}{\lambda_{ct}}$ is the stochastic discount factor of the households, which are described next.

2.2 Households

The representative household's problem is

$$\max_{C_t, D_t, E^s_t, E^r_t} E\sum_{t=0}^{\infty}\beta^t\left[\frac{\left(C_t - \kappa C_{t-1}\right)^{1-\varsigma_c} - 1}{1-\varsigma_c} + \varsigma_0\frac{D_t^{1-\varsigma_d} - 1}{1-\varsigma_d}\right], \tag{14}$$

subject to

$$
\begin{aligned}
C_t + D_t + E_t^s + E_t^r &= W_t + R_{t-1}^d D_{t-1} + R_t^{e,s} E_{t-1}^s + R_t^{e,r} E_{t-1}^r - T_t, \qquad (15)\\
E_t^s &\geq 0,\\
E_t^r &\geq 0.
\end{aligned}
$$

Households value consumption, C_t, and value the liquidity services of bank deposits, D_t; $0 < \kappa < 1$ is the habit persistence parameter, $\varsigma_c > 0$ captures the intertemporal elasticity of substitution, $\varsigma_0 > 0$ governs the utility weight on deposits relative to consumption, and $\varsigma_d > 0$ is the inverse elasticity of households demand for deposits with respect to changes in the interest rate. We put deposits in the utility function in lieu of modeling a particular transactions technology. And for simplicity, we assume that households supply labor inelastically and have normalized the supply of labor to be one.[3] Household assets include deposits, D_t, which pay a gross real rate R_t^d, and two types of bank equity: E_t^s is equity in a "safe" bank, which lends to a safe firm and pays $R_{t+1}^{e,s}$ next period; E_t^r is equity in a "risky" bank, which lends to a risky firm and pays $R_{t+1}^{e,r}$. The returns on equity are of course not known when the household invests. By contrast, the return on deposits is known, and deposits are protected by deposit insurance; deposits are the safe asset in our model. Finally, households pay lump sum taxes, T_t, to fund the government's deposit insurance program.

The household's first order conditions include:

$$
\begin{aligned}
C: &\quad (C_t - \kappa C_{t-1})^{-\varsigma_c} - \beta\kappa E_t (C_{t+1} - \kappa C_t)^{-\varsigma_c} - \lambda_{ct} = 0, \qquad (16)\\
D: &\quad \varsigma_0 D_t^{-\varsigma_d} - \lambda_{ct} + E_t \beta \lambda_{ct+1} R_t^d = 0, \qquad (17)\\
E^s: &\quad -\lambda_{ct} + E_t \beta \lambda_{ct+1} R_{t+1}^{e,s} + \zeta_t^s = 0, \qquad (18)\\
E^r: &\quad -\lambda_{ct} + E_t \beta \lambda_{ct+1} R_{t+1}^{e,r} + \zeta_t^r = 0, \qquad (19)
\end{aligned}
$$

where λ_{ct}, ζ_t^s and ζ_t^r are the Lagrangian multipliers for the budget constraint and the two non-negativity constraints.

If households did not value deposits for their liquidity services ($\varsigma_0 = 0$), (17) would be the standard RBC Euler equation, and R_t^d would be the standard CAPM rate. But households do value deposits in our model, and R_t^d is below the CAPM rate. Equity is not a safe asset, and it does not provide liquidity services. So, deposits will be the cheaper source of funding for banks. This fact will play an important role in what follows.

[3] While the total supply of labor is fixed, its distribution across safe and risky firms is market determined.

2.3 Banks

Banks are at the heart of our model. First, we set the stage by describing their incentives to take excessive risk. Second, we discuss the banking sector in some detail.

2.3.1 Incentives to Take Excessive Risk and Capital Requirements

We saw from the section on firms that $E_t R^r_{t+1} < E_t R^s_{t+1}$. So, why would a profit maximizing bank ever invest in a risky firm? Limited liability and government deposit insurance are the culprits here. Limited liability shields the bank from downside risk. Moreover, deposit insurance actually subsidizes risk taking; it makes bank deposits the safe asset, lowering the cost of issuing deposits, and allowing the bank to expand its portfolio of safe or risky loans. In what follows, we will see that if the expected return on investment in a safe firm falls, due say to a negative TFP shock, the bank may be tempted to take a flier on the risky firm.

As we will see, capital requirements are a potential remedy for excessive risk taking. In what follows, we will consider a requirement that says equity finance cannot fall below a fraction γ_t of the bank's loans. A high γ_t requires the bank and its equity holders to keep more skin in the game, and it shrinks the bank's portfolio since equity finance is more expensive than deposit finance.

2.3.2 The Banking Sector

A measure one continuum of perfectly competitive banks start operating each period, and they live for two periods. In the first period, a bank issues equity and deposits to households, and uses the proceeds to make loans to firms; in the second period, the bank receives the return on its investments and liquidates its assets and liabilities.

More specifically, in period t, the bank creates a loan portfolio by directing a fraction σ_t of its loans to a risky firm; the remainder of its loans go to a safe firm.[4] Since $R^r_{t+1} = R^s_{t+1} + \frac{\varepsilon_{t+1}}{Q_t}$, the ex-post return on the portfolio will be $R^s_{t+1} + \sigma_t \frac{\varepsilon_{t+1}}{Q_t}$. Note that $nw_{t+1} \equiv \left(R^s_{t+1} + \sigma_t \frac{\varepsilon_{t+1}}{Q_t} \right) l_t - R^d_t d_t$ is the bank's net worth in period $t+1$, where l_t is the total amount of loans. If nw_{t+1} is positive, the bank pays its depositors and distributes the rest to its equity holders. If it is negative, the bank declares bankruptcy; its depositors are protected by deposit insurance, but its equity holders get nothing.

The bank's objective is to maximize the expected return of its equity holders, whose stochastic discount factor is $\psi_{t,t+i}$. Let ε^*_{t+1} be the realization of the idiosyncratic shock

[4] Our assumption that a bank only deals with one safe and one risky firm comes at no loss of generality because all the safe firms are identical, and diversification among the risky firms does not take full advantage of the bank's limited liability. See Collard et al (2017) for a more formal exposition of this result.

below which the bank's net worth is negative; that is, $\left(R^s_{t+1} + \sigma_t \frac{\varepsilon^*_{t+1}}{Q_t}\right) l_t - R^d_t d_t = 0$. Since the distributions of aggregate and idiosyncratic shocks are independent of each other, we can nest expectations with respect to the idiosyncratic shock within the expectation of the aggregate and idiosyncratic shocks, and the representative bank's maximization problem can be written as:

$$\max_{l_t, d_t, e_t, \sigma_t} E_t \left\{ \psi_{t,t+i} \left[\int_{\varepsilon^*_{t+1}}^{\infty} nw_{t+1} \, \mathrm{d}G(\varepsilon_{t+1}) \right] \right\} - e_t \tag{20}$$

subject to

$$\begin{aligned} & l_t = e_t + d_t \\ & e_t \geq \gamma_t l_t \\ & l_t \geq 0 \\ & \underline{\sigma} \leq \sigma_t \leq \bar{\sigma} \end{aligned} \tag{21}$$

where e_t is equity issued to households. The first constraint is the bank's balance sheet, and the second is the bank's capital requirement. The third constraint rules out short selling; its role will be discussed in Section 4. The fourth imposes limits on the fraction of a bank's portfolio that can go to safe or risky loans. In our calibrations, $\bar{\sigma}$ is set equal to 0.99 and $\underline{\sigma}$ is set equal to 0.01; so, banks can get very close to totally safe or totally risky portfolios if they so choose.[5]

The bank's first order conditions can be found in Appendix A.1. They are not particularly elucidating. In the next subsection, we discuss the bank's basic tradeoff when it decides how risky to make its portfolio of loans.

2.3.3 The Bank's Dividends, and Its Choice of σ_t.

In Appendix A.5, we derive the bank's expected (discounted) dividend function,

$$\Omega(\sigma_t;\, l_t,\, d_t,\, e_t) = E_t \left[\psi_{t,t+i} l_t \left(\omega_1 + \omega_2 \right) \right], \tag{22}$$

[5] These limits on σ_t are necessary for the numerical methods that follow.

where

$$\omega_1 \equiv \left(R^s_{t+1} - R^d_t (1-\gamma_t) - \frac{\xi \sigma_t}{Q_t} \right) \left(1 - G(\varepsilon^*_{t+1})\right) \tag{23}$$

$$\omega_2 \equiv \left(\frac{\sigma_t}{Q_t} \right) \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\varepsilon^*_{t+1}+\xi}{\tau\sqrt{2}} \right)^2} \tag{24}$$

and where $1 - G(\varepsilon^*_{t+1})$ is the probability that the bank will not default.

The first component, ω_1, is the return on a loan portfolio with a fraction σ_t going to a risky firm; $-\xi$ is the (negative) expected value of the idiosyncratic shock. The second component, ω_2, is a bonus attributable to the bank's limited liability; the higher is the standard deviation of the idiosyncratic shock, τ, the higher is the upside potential for a risky loan, while the downside risk is protected by limited liability.

Increasing σ_t makes the portfolio more risky. More risk decreases the ex-post return on the bank's portfolio, but it increases the bonus from limited liability. This is the tradeoff that a bank faces.

2.4 The Government

The government provides deposit insurance, and collects taxes to pay for it. Given the Ricardian nature of the model, a lump sum tax, T_t, can balance the budget each period without distorting private decision making. In Appendix C.1, we show the tax necessary to support the insurance scheme is

$$T_t = \frac{\sigma_{t-1} L_{t-1}}{Q_{t-1}} \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{R^d_{t-1}(1-\gamma_{t-1})Q_{t-1} - R^s_t Q_{t-1} + \xi \sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau} \right)^2} - \tag{25}$$
$$\frac{1}{2} \left(R^s_t L_{t-1} - \frac{\sigma_{t-1}\xi}{Q_{t-1}} L_{t-1} - R^d_{t-1} D_{t-1} \right) \left[1 + \operatorname{erf}\left(\frac{R^d_{t-1}(1-\gamma_{t-1})Q_{t-1} - R^s_t Q_{t-1} + \xi \sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau} \right) \right],$$

where L_t is the aggregate amount of loans provided by the banking sector. As might be expected, more risk taking (a higher σ_{t-1}) and/or a higher variance (τ) of the idiosyncratic shock increases the taxes required to protect deposits.

2.5 Analytical Characterization of Equilibrium

We are able to derive some analytical results that enhance our understanding of the model's equilibrium, and how to calculate it. More generally, we will require numerical methods.

2.5.1 Two Propositions and a Corollary

As discussed in the section on households, deposits are a cheaper source of bank funding than equity. So, a bank will fund as much of its loans by issuing deposits as is allowed by the capital requirements. We formalize this argument and prove the following proposition in Appendix D.

Proposition 1. *In equilibrium, capital requirements always bind; that is,* $e_t = \gamma_t l_t$*.*

The next proposition, and its corollary, show that we need only consider two values of the bank's portfolio risk parameter, σ_t, when we derive the model's equilibrium. The proposition is established in Appendix D.

Proposition 2. *The expected dividends function of banks,* $\Omega(\sigma_t;\, l_t,\, d_t,\, e_t)$*, is convex in* σ_t*. This result holds for arbitrary (and not necessarily continuous) distributions of the idiosyncratic shock.*

Corollary. There are no equilibria with $\underline{\sigma} < \sigma_t < \bar{\sigma}$.

The intuition for this proposition and its corollary is as follows: If σ_t is high enough, the bank will be bankrupt for low values of ε_t anyway, so it might as well take on as much risk as possible to maximize the portfolio's upside potential for high values of ε_t. If σ_t is low enough, the bank will not be bankrupt even for low values of ε_t, and the value of limited liability is negated; the bank might as well take on the minimum risk to raise the expected value of its portfolio.

2.5.2 Equilibrium and Aggregation

We consider a competitive equilibrium in which each bank takes aggregate prices as given. Appendix E lists all the equilibrium conditions of our model. In this subsection, we only present the equilibrium conditions that are not already included in the preceding subsections. We let μ_t denote the fraction of banks with risky portfolios (banks that choose $\sigma_t = \bar{\sigma}$) at date t; the remaining fraction $1 - \mu_t$ are safe banks ($\sigma_t = \underline{\sigma}$).

The fraction μ_t is endogenously determined by equity positions of households: we have $\mu_t = \frac{E_t^r}{E_t^r + E_t^s}$. At any point in time, the economy may be in a safe equilibrium (with $\mu_t = 0$), a risky equilibrium (with $\mu_t = 1$), or a mixed equilibrium (with $0 < \mu_t < 1$).

Each bank within a group (safe or risky) is alike and solves the same maximization problem in which it chooses l_t^i, d_t^i, e_t^i according to its type $i \in \{s, r\}$. The aggregate loans to the (representative) safe firm come from two sources: 1) from all safe banks (of measure $1-\mu_t$) that allocate $1 - \underline{\sigma}$ share of their loan portfolio to safe projects and 2) from all risky banks (of measure μ_t) that allocate $1-\bar{\sigma}$ share of their loan portfolio to safe projects. Therefore, the equilibrium restrictions linking our bank-level and firm-level variables representing loans are

$$Q_t K^s_{t+1} = (1-\underline{\sigma})(1-\mu_t) l^s_t + (1-\bar{\sigma}) \mu_t l^r_t. \tag{26}$$

Similarly,

$$Q_t K^r_{t+1} = \underline{\sigma}(1-\mu_t) l^s_t + \bar{\sigma} \mu_t l^r_t. \tag{27}$$

The aggregate bank loans are linked to the individual bank loans by: $L^r_t = \mu_t l^r_t$ and $L^s_t = (1-\mu_t) l^s_t$. Therefore, we can describe the latter two equations by using aggregate loans

$$Q_t K^s_{t+1} = (1-\underline{\sigma}) L^s_t + (1-\bar{\sigma}) L^r_t, \tag{28}$$

$$Q_t K^r_{t+1} = \underline{\sigma} L^s_t + \bar{\sigma} L^r_t. \tag{29}$$

The equity positions taken by households, in turn, determine the equity positions of individual banks: $E^r_t = \mu_t e^r_t$ and $E^s_t = (1-\mu_t) e^s_t$. The returns on the equity positions taken by households at date t are linked to the dividends paid by banks at date $t+1$. We have:

$$E^r_t R^{e,r}_{t+1} = (\omega^r_1 + \omega^r_2) L^r_t, \tag{30}$$

$$E^s_t R^{e,s}_{t+1} = (\omega^s_1 + \omega^s_2) L^s_t, \tag{31}$$

where we use that $\max\left[nw^r_{t+1}, 0\right]$ is linear in loans; ω_1 and ω_2 were defined in equations (23) and (24). Deposits held by households are issued by (safe and risky) banks: $D_t = D^s_t + D^r_t$ where $D^s_t = L^s_t - E^s_t$ and $D^r_t = L^r_t - E^r_t$.

The equilibrium restrictions linking our aggregate and individual firm-specific variables are straightforward, but cumbersome in terms of notation. We state the restrictions in Appendix E. The market-clearing conditions for labor, capital, and goods are

$$H^s_t + H^r_t = 1, \tag{32}$$

$$K^s_t + K^r_t = K_t, \tag{33}$$

and

$$Y^s_t + Y^r_t = C_t + I^g_t. \tag{34}$$

3 Calibration and Steady-State Capital Requirements

Our calibrated parameters are reported in Table 1. We use standard values for the discount factor β, the capital share α, the intertemporal elasticity of substitution ϱ_c, and the depreciation rate δ.

We consider loans to be risky if they are made to firms with a debt-to-EBITDA ratio above 6 in the leveraged loan market.[6] We choose τ, the standard deviation of the risky firm's idiosyncratic shock, to match the variance of returns on a risky project in our model to the variance of returns from lending to a firm with a debt-to-EBITDA ratio of 6. In each case, we focus on variances conditional on starting from the non-stochastic steady state of our model. Appendix G provides the details of our procedure. Given τ, we fix the value of ξ, the average penalty from financing risky projects, so that a 10% steady- state capital requirement prevents lending to risky firms. We note that our choice of 10% is consistent with the static values of capital requirements proposed by Basel III; it also lies within a span of values usually considered in the literature on optimal capital regulation.

Why do we not try to calculate an optimal steady-state capital requirement? We show in Appendix H that alternative choices of τ and ξ would support a wide range of steady-state capital requirements. This suggests that a model like ours is not suitable for any attempt to pin down the optimal steady-state value.

To match the data on interest rate spreads, we introduce costs of banking in our quantitative model. We assume that these costs are linked to the provision of loans. In particular, each period the bank incurs an additional cost, fl_t, that is paid out of its current profits.[7] And when a bank defaults, the household has to pay a higher tax to the deposit insurance fund to cover this cost of banking. Appendix C.2 and Appendix A.6 provide further details on the implications of this cost for the lump sum tax, T_t, and on the first order conditions for the optimization problem of banks. We choose f to make the average spread between the safe loan rate and the deposit rate equal to 2.26 percent per annum; we take this value from Collard et al. (2017). The parameter ς_0 measures the utility of deposits in the steady state. We set the value of ς_0 to make the interest rate on bank deposits equal to 0.86% per quarter, a value we borrow from an estimate in Begenau (2019). Finally, our setup for investment adjustment costs mimics the one used by Altig et al. (2011). We pick the value of ϕ consistent with the broad range from their analysis and related literature.

[6] EBITDA is earnings before interest, taxes, depreciation, and amortization.

[7] Allowing for these banking costs serves to calibrate the steady-state equilibrium of our model; but it has no effect on the equilibrium dynamics we discuss in subsequent sections, because labor supply is inelastic in our model. For this reason, we have suppressed this factor in the equations above.

4 Numerical Methods

Since our model involves occasionally binding nonnegativity constraints on bank loans, we need to rely on nonlinear solution methods. We apply the Occbin toolkit developed in Guerrieri and Iacoviello (2015). This solution algorithm modifies a first order perturbation method and employs a guess-and-verify approach to obtain a piecewise linear solution.[8] The solution reflects the endogenous transition between safe and risky regimes, depending on the size of a shock and the state vector, and thus it is highly nonlinear. The algorithm has advantages over nonlinear projection methods because it is computationally fast and can be applied to nonlinear models with a large number of state variables, such as ours.

So why did we complicate matters by imposing nonnegativity constraints on loans? We needed to rule out the short-selling of assets (or negative loans). To see why, suppose banks are in the safe equilibrium; in this case, risky loans are overpriced compared to safe loans (because expected returns on risky loans are relatively lower in the safe equilibrium); absent short-selling restrictions, each bank would want to short risky loans. Similar reasoning applies to the risky equilibrium, in which the banks in our model would short safe loans. In either of these cases, arbitrageurs would force the expected returns on safe and risky loans to equality. And this would result in the mixed equilibrium (described in Section 2.5.2) in which $0 < \mu_t < 1$.

5 The Ramsey Policy and Its Numerical Derivation

To compute optimal capital requirements, we focus on the Ramsey problem, conditional on the restrictions of the decentralized equilibrium. The Ramsey program selects the path of capital requirements that maximizes the conditional expectation of the household's utility as of time zero. More precisely, following a dual approach, the Ramsey planner chooses the sequence of capital requirements $\{\gamma_t^*\}_{t=0}^{\infty}$ to maximize the household utility function, (14), subject to the equilibrium conditions implied by the optimality conditions of households, firms and banks, and the market clearing conditions. The non-negativity and short-selling restrictions that we noted above complicate this Ramsey problem. We proceed by proposing a natural candidate for the solution and then verifying that the proposed solution does indeed maximize the objective function, (14).

Our proposed solution is to consider the sequence of capital requirements $\{\gamma_t^*\}_{t=0}^{\infty}$ that is set at the lowest level necessary to prevent risk taking – given the realizations of the shocks – at any date t. This sequence dominates any alternative path $\left\{\gamma_t^A\right\}_{t=0}^{\infty}$ in which $\gamma_t^A = \gamma_t^*$

[8] See Guerrieri and Iacoviello (2015) for a discussion of the accuracy of this type of solution method.

for $t \neq t_k$ and $\gamma_t^A = \gamma_t^* + \Delta$ for $t = t_k$ and some $\Delta \neq 0$. 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 does not depend on the size of $\Delta > 0$. When $\Delta < 0$, banks switch to funding socially inefficient risky projects in period t_k under $\left\{\gamma_t^A\right\}_{t=0}^{\infty}$. The decrease in the capital requirement involves an output loss of ξK from making risky loans, but it may increase the liquidity services that enter into household utility. The trade-off between these two considerations determines the impact on welfare. For a small decrease in capital requirements (i.e. negative values of Δ close to zero), the former consideration is more important. Why? Since banks jump to the risky equilibrium, the lower capital requirement entails a discrete drop in welfare, arising from the drop in output. By contrast, the welfare gain (or loss) associated with liquidity provision is a second order change.

Our reasoning above establishes that the Ramsey planner's objective function has a local maximum along the path $\{\gamma_t^*\}_{t=0}^{\infty}$. To show that this is indeed a global maximum, we must check the welfare effect of a large decrease in capital requirements; in this case, liquidity considerations will not be of second order. To see how liquidity considerations compare to the welfare loss associated with inefficient risk taking, we compare (numerically) the welfare measure under our candidate for optimal policy to welfare under an alternative policy that maximizes the benefit of liquidity provision under the risk-taking regime. All the equilibria under the risk-taking regime have the same level of expected output; so, we only need to consider the policy that maximizes liquidity provision. The gains from liquidity services are maximized when $\gamma_{t_k}^A = 0$. Therefore, we need to compare conditional welfare under $\{\gamma_t^*\}_{t=0}^{\infty}$ to the alternatives that let the capital requirement go down to zero, in some periods.

To check quantitatively if setting capital requirements to zero becomes optimal in response to shocks, we use a variant of the OccBin algorithm. We consider a horizon J and construct all possible combinations of periods from 1 to J in which capital requirements are hardwired to go to zero whenever a switch to the risk-taking regime is made, but are set at $\{\gamma_t^*\}_{t=0}^{\infty}$ otherwise. Then, for each combination, we calculate the conditional welfare and compare it against the conditional welfare of keeping capital requirements at $\{\gamma_t^*\}_{t=0}^{\infty}$. We verify that the proposed path of $\{\gamma_t^*\}_{t=0}^{\infty}$ that makes capital requirements just large enough to prevent excessive risk-taking incentives is, in fact, globally optimal in our parameterization.[9]

[9] We did verify that, in response to shocks, the optimal capital requirements are 0 over some periods if the weight on deposits in the utility function is high enough. However, this higher weight on deposits renders the steady-state spread between the risk-free interest rate and interest rate on deposits counterfactually large.

6 Optimal Dynamic Capital Requirements

In this section, we show how a Ramsey Planner would set capital requirement ratios, γ_t, in response to various shocks that can cause excessive risk taking. All of the shocks we consider in this section follow exogenously set AR(1) processes, which are specified below.[10] We take two steps in preparation for our discussion here. First, we ask what might trigger a risk-taking episode in the first place. And second, we show how exogenous shocks to the Planner's policy instrument – capital requirements – would affect financing decisions and real allocations.[11]

6.1 What Triggers an Excessive Risk-Taking Episode?

The answer to this question is rather complex because the banker's maximization problem has so many moving parts. We give a detailed answer in Appendix F; here we offer a simpler explanation that focuses on the main forces at work.

Consider the expected dividends for safe and risky firms, $\Omega_t^s \equiv \mathit{\Omega}(\underline{\sigma};\, l_t,\, d_t,\, e_t)$ and $\Omega_t^r \equiv \mathit{\Omega}(\bar{\sigma};\, l_t,\, d_t,\, e_t)$ respectively. Anything that would make $\Omega_t^r - \Omega_t^s$ go positive will trigger a risk-taking episode. Equation (22) specifies $\mathit{\Omega}(\sigma_t;\, l_t,\, d_t,\, e_t)$ for all values of σ_t, where it will be recalled that

$$\varepsilon_{t+1}^{*} = -\frac{Q_t}{\sigma_t}\left[R_{t+1}^{s} - R_t^{d}\left(1-\gamma_t\right)\right] \tag{35}$$

is the realization of a bank's idiosyncratic shock below which its net worth is negative, and $G(\varepsilon_{t+1}^{*})$ is the probability that the bank will fail. Implicit in the formulation of the banker's problem, (20), is the fact that $G'(\varepsilon_{t+1}^{*}) > 0$ and $G(\varepsilon_{t+1}^{*}) \to 0$ as $\varepsilon_{t+1}^{*} \to -\infty$.

For purely expositional purposes, we will in this subsection suppose that $\underline{\sigma} = 0$ and $\bar{\sigma} = 1$. With these simplifications, (22) implies

$$\Omega_t^s = E_t\left[\psi_t l_t\left(R_{t+1}^{s} - R_t^{d}\left(1-\gamma_t\right)\right)\right] \text{ and} \tag{36}$$

$$\Omega_t^r = E_t\left[\psi_t l_t\left(\left(R_{t+1}^{s} - R_t^{d}\left(1-\gamma_t\right) - \frac{\xi}{Q_t}\right)\left(1 - G(\varepsilon_{t+1}^{*})\right) + \frac{\tau}{Q_t\sqrt{2\pi}} e^{-\left(\frac{\varepsilon_{t+1}^{*}+\xi}{\tau\sqrt{2}}\right)^2}\right)\right], \tag{37}$$

where $\psi_t \equiv \beta\frac{\lambda_{ct+1}}{\lambda_{ct}}$ is the household's stochastic discount factor, and where it will be recalled that

[10] The rest of the parameter settings are given in 1, except that here we set $\varphi = 100$ and $\kappa = 0$.

[11] For the purposes of this section, we have set the steady-state capital requirement at 10.1 percent, 0.1 percent higher than is necessary to avoid excessive risk taking in the steady state. This facilitates our numerical solution methods.

$$R^s_{t+1} = \alpha \left\{ \frac{A_{t+1}}{Q_t} \left(\frac{H^s_{t+1}}{K^s_{t+1}} \right)^{1-\alpha} + (1-\delta) \frac{Q_{t+1}}{Q_t} \right\}. \tag{38}$$

What might turn $\Omega^r_t - \Omega^s_t$ positive, triggering a risk-taking episode? The obvious culprit is the interest rate spread $R^s_{t+1} - R^d_t (1-\gamma_t)$. An expected narrowing of this spread will decrease Ω^s_t more than Ω^r_t since $1 - G(\varepsilon^*_{t+1})$ is less than one in the risk-taking regime. Moreover, a narrowing of the spread has a secondary effect on Ω^r_t that is a little more subtle: (35) implies that ε^*_{t+1} will rise. The presence of ε^*_{t+1} (instead of $-\infty$) in the bank's expected profits, (20), represents the value of limited liability to banks. Idiosyncratic shocks below this cut-off point cannot lower the bank's profits. An increase in ε^*_{t+1} would enhance the value of the shield of limited liability and increase Ω^r_t.[12] Note finally that if a risk taking episode is triggered, there will be a jump in σ, and therefore a further jump in ε^*_{t+1} .

So, what might narrow the interest rate spread and provoke a risk-taking episode? There are a number of possibilities. Perhaps the most obvious would be a fall in the expected return on safe assets; for example, an expected fall in TFP could trigger a risk-taking episode. Two parameters in (37) are also of interest. An increase in the standard deviation of the idiosyncratic shock, τ, will raise Ω^r_t since it increases the upside potential of the risky asset (while the downside potential is unchanged because of limited liability). The second parameter is the expected value of the risky firm's idiosyncratic shock is $-\xi$; ξ is the average penalty for investing in the risky asset. A fall in this parameter would also raise Ω^r_t.

Note also that a loosening of the capital requirement, γ_t, would decrease the interest rate spread and could trigger a risk-taking episode. A loosening of the capital requirement allows the bank to fund more of its loans with deposits; this reduces the cost of banking and allows the bank to keep less skin in the game. The bank expands its lending and switches to risky loans. And note finally that a dynamic capital requirement could hold $\Omega^r_t - \Omega^s_t$ constant at it's steady-state value; banks would never leave the safe equilibrium. As seen in Section 5, this option is the Ramsey Planner's policy.

The intuitive exposition just given relied upon two simplifying assumptions – one made explicit, and the other implicit – that must now be undone. The explicit assumption was that $\underline{\sigma} = 0$ and $\bar{\sigma} = 1$. In the numerical analysis that follows, $\underline{\sigma}$ is set equal to 0.01 and $\bar{\sigma}$ is set equal to 0.99; in equilibrium, there must be both safe and risky loans (and firms). The implicit assumption was that a bank could observe both Ω^r_t and Ω^s_t, and then choose its loan portfolio accordingly. But, we cannot have both Ω^r_t and Ω^s_t in equilibrium. If we are

[12] It is hard to see these results in (37) without investigating a number of special cases, some involving the absolute value of $\varepsilon^*_{t+1} + \xi$. These special cases are relegated to Appendix F.

in a safe equilibrium, we have Ω_t^s, and Ω_t^r is an off-equilibrium object; during a risk taking episode, we have Ω_t^r, and Ω_t^s is an off-equilibrium object.

However, there is an equilibrium interest rate spread – whose evolution is closely related to $\Omega_t^r - \Omega_t^s$ – that we can track:

$$S_t \equiv E_t \left[R_{t+1}^{e,r} - R_{t+1}^{e,s} \right]. \tag{39}$$

S_t is the expected spread between the returns on risky and safe equity. Because of our minimum scale assumptions, a small amount of risky loans will be extended in the safe regime, and conversely, a small amount safe loans will be extended in the risky regime; so, the returns on equity are equilibrium objects. In a risk-taking episode, S_t turns positive. Once the episode is over, the spread turns negative.[13]

6.2 Capital Requirement Shocks

The next two sections illustrate the transmission mechanism for capital requirement policy. And in particular, we show that increases and decreases in capital requirements have asymmetric effects on bank decision making and economic outcomes.

6.2.1 An Increase in Capital Requirements

Figure 1 shows the effects of a one percent increase in the capital requirement ratio, γ_t; this shock has a persistence parameter of 0.9. An increase in the capital requirement forces a bank to shift its funding mix from deposits to equity; this shift increases the cost of funding a given amount of loans since deposits have liquidity value, and they will be held by the households at a lower rate of return. The shock does make the bank safer by requiring it to keep more skin in the game.

Note that the Modigliani-Miller Theorem does not hold in our model, since once again deposits are valued for their transactions services. So, even though the economy stays in a safe equilibrium, tighter capital requirements can have real effects on the macroeconomy.

[13]There is a simple relationship between S_t and $\Omega_t^r - \Omega_t^s$ when computing Ω_t^r and Ω_t^s conditional on, respectively, the risky and safe loans actually extended (rather than the *desired* amount of loans). In that case, $S_t \equiv E_t \left[R_{t+1}^{e,r} - R_{t+1}^{e,s} \right] = E_t \frac{\Omega_{t+1}^r}{E_t^r} - E_t \frac{\Omega_{t+1}^s}{E_t^s}$. The thought experiment by which a banker compares the expected dividends for a desired level of loans is intuitive, but we solve the model by referring to the Lagrange multipliers on the non-negativity constraints for safe and risky loans. When extending safe loans leads to higher expected dividends, a banker would want to short-sell risky loans, turning the corresponding Lagrange multiplier positive; analogously, when extending risky loans leads to higher expected dividends, a banker would want to short-sell safe loans. These two conditions allow us to determine which regime applies in any period more easily than attempting to construct $E_t \frac{\Omega_{t+1}^r}{E_t^r}$ and $E_t \frac{\Omega_{t+1}^s}{E_t^s}$, whose computation requires taking a stand on the entire path of future actions.

More precisely, an increase in the capital requirement acts like a tax hike on banks. Households, who own the banks, are effectively poorer. They cut back on consumption, and since labor is inelastically supplied, their savings increase correspondingly. But under our calibration, the movements in consumption, investment and output are tiny, as can be seen in Figure 1. The real side of the economy is hardly affected.

There are first order effects in the financial sector, and they can affect household utility. First and foremost, the increase in equity funding reduces the bank's demand for deposits, and the deposit rate falls. Moreover, the increase in household savings pushes up the supply of deposits, which reinforces the decrease in the deposit rate. Deposits make up close to 90 percent of bank funding in our calibration. Somewhat paradoxically, the increase in capital requirements, and the subsequent fall in the deposit rate, end up *reducing* the cost of banking.[14] However, the large drop in deposits, coupled with the (almost imperceptible) fall in consumption, decreases household utility, as can be seen in the last panel in Figure 1.[15]

Over time, these movements reverse themselves. The capital requirement falls, and deposits recover. The capital stock falls, increasing the marginal product of capital and R^s, which pushes Ω^s up relative to Ω^r. The economy reverts to its steady state.

6.2.2 A Decrease in Capital Requirements:

The dashed lines in Figure 2 show the response to a 1 percent decrease in the capital requirement, with an auto-regressive coefficient of 0.9. Deposits rise and bank equity falls, as the lower capital requirement allows banks to switch to the cheaper source of funding. As explained in Section 6.1, a loosening of the capital requirement immediately triggers a risk-taking episode. On average, risky firms produce less output since a risky firm's idiosyncratic shock has a negative expected value; so, output and income fall substantially.[16] Consumption and investment also fall. In subsequent periods, the demand for capital falls, as does its price, Q_t. The fall in Q_t, coupled with the jump in σ_t, increases the cut-off point ε^*_{t+1} discussed in Section 6.1, making risky loans more attractive; Ω^r_t and $R^{e,r}_{t+1}$ rise. The spread S_t immediately goes positive. These events are pictured in Panels 5 and 7.

Over time, the capital requirement rises and the process described above reverses itself. When S_t falls to zero, σ_t jumps back to its lower bound, and the economy jumps back to

[14] Begenau (2019) also finds that an increase in capital requirements can reduce the cost of bank funding and increase lending.

[15] Welfare is calculated as the present discounted value of utility at a given point in time; is moves as the state variable change.

[16] Put another way, some of the risky loans fail, destroying bank equity and increasing the taxes necessary to insure deposits. So, output and income fall.

a safe equilibrium. Capital is more productive in a safe equilibrium, since lending to the inefficient risky firms is almost eliminated. This creates a jump in the price of capital, Q_t,and a jump in the return on safe loans, as can be seen in (38); the expected return on safe equity spikes. Gradually, the economy returns to its steady state.

Takeaways:

Positive and negative shocks to the capital requirement have asymmetric effects on the economy, and they are not the mirror images found in linear models. Loosening capital requirements triggers an excessive risk-taking episode, and consumption and output fall. For comparison, the solid lines in Figure 2, repeat the responses shown in Figure 1; the responses of consumption and output are so small as to be imperceptible with the re-scaling of the axes. Loosening capital requirements produces a major disruption on the real side of the economy; for a tightening of capital requirements, what happens in the financial sector stays in the financial sector.

6.3 TFP Shock

TFP shocks have played a major role in RBC modeling. Figure 3 illustrates the effects of a contractionary TFP shock; A_t falls by 1.5 percent (or one standard deviation), and has a persistence parameter of 0.95. In each panel, the dashed line shows what would happen if γ_t were to be held constant at its steady-state value; the solid line shows what would happen if the Ramsey Planner set the path of γ_t.

We begin with the case of fixed capital requirements. Since the shock is auto-correlated, today's TFP shock lowers the expected marginal productivity of capital for the next period, and thus the expected return on safe assets. As explained in Section 6.1, this triggers a risk-taking episode. $R^{e,s}_{t+1}$ falls and the spread S_t jumps positive. Risky firms produce less output on average; so, output and income fall substantially, as does consumption. As output and the marginal productivity of capital fall, the demand for capital falls, lowering the investment price, Q_t. For use in Section 7 below, we also track the credit-to-GDP ratio. It falls, as under our calibration, bank loans decrease more quickly than GDP.

Over time, the TFP shock dissipates and the process described above reverses itself. Among other things, the falling capital stock raises the marginal productivity of capital and the return on safe assets, and also the price of investment. S_t falls, and jumps negative after σ_t drops to its lower bound, and the economy jumps back to a safe equilibrium. The credit-to-GDP ratio rises, and then midway starts to fall.

Next, we turn to the Ramsey Planner's solution, shown by the solid lines in Figure 3. The Planner's policy is to set capital requirements just tight enough to keep safe loans attractive;

as we have seen, any higher would unnecessarily deprive households of the deposits that they value. γ_t jumps on impact, and falls back to its steady-state value as the TFP shock dissipates.

While the Planner's policy avoids risk-taking episodes, it cannot undo the damage done by the TFP shock itself. The shock lowers the household's net worth, and they respond by decreasing consumption and increasing savings/investment. All this is familiar from the RBC literature. Indeed, absent the possibility of excessive risk taking, our model has no banking frictions; in essence, it reduces to the standard RBC model in which there is no role for macroeconomic policy. It may be interesting to note that the gap between the paths of consumption in the third panel is largely determined by the size ξ, the expected loss on risky loans; ξ is a measure of the economic inefficiency in our model.

Takeaways: A one standard deviation shock to TFP causes a 1.5 percent decrease in output. However, the optimal capital requirement needs only a modest adjustment, an increase from 10 percent to 10.15 percent. Note also that after a few quarters, the path of the investment price, Q_t, in the inefficient solution (inversely) tracks the path in the Planner's solution rather closely. After its initial fall, the credit-to-GDP ratio rises and then falls midway through the cycle; optimal capital requirements do not follow the prescription laid out by the Bases III accords.

6.4 An Expansionary Investment Technology Shock.

Here we study a positive η_t shock in the equation for net investment, (11). The shock has a persistence parameter of 0.8, and we calibrate the size of the shock to increase output by 1% at its peak, roughly on a par with the TFP shock described previously. Figure 4 illustrates the effects of this shock. Once again, the dashed lines show what would happen if the capital requirement were kept at its steady-state value, while the solid lines represent the Ramsey solution.

This shock was not considered in Section 6.1, but its effects are readily translatable to the discussion there. A positive shock to investment in period t increases the supply of capital next period, K_{t+1}, lowering the expected marginal product of capital and the expected return on the safe asset. The expected return on safe equity falls, and a risk-taking episode is begun, even though the shock itself is expansionary.

Note that the expected return on safe equity only drops for one period. To see why, note that the decrease in the marginal product of capital causes the price of capital, Q_{t+1}, to fall, and this raises the return on safe loans in period $t+2$. However, the damage is already done; the risk-taking episode has already been triggered, as documented by the jump in S_t. The risky firms produce less output on average, and output and consumption fall. From here on, the story is much the same as before. The investment shock decays over time and the process gradually reverses itself. Note that there is an upward spike in the expected return on safe loans when the economy jumps back to a safe equilibrium.

The solid lines illustrate what would happen if the Ramsey Planner set the path of γ_t. The Planner raises the capital requirement just enough to offset the switch to excessive risk taking. Consumption and investment rise more in this case since there are no bankruptcies and equity losses to lower household income.

Takeaways: In this example, the Planner raises capital requirements as the economy goes into a boom period, which may be thought to be in line with Basel III's cyclical buffers; however the capital-to-GDP ratio falls initially. Once again, this ratio subsequently rises, and then falls midway through the cycle. The optimal adjustment in the capital requirement is again small; γ_t only rises from 10% to a little over 10.2%. But the distance between the solid and dashed lines is substantial. Note that the path followed by the investment price, Q_t, in the inefficient solution (inversely) tracks the Planner's solution, but rather loosely.

6.5 A Volatility Shock

In the steady state, the standard deviation of the idiosyncratic shock, τ, affecting risky firms is 5.5%. Our volatility shock increases the standard deviation by 15 basis points, after which it follows an AR(1) process (with persistence parameter 0.8) back to 5.5%. As explained in Section 6.1, an increase in volatility raises the expected return on risky loans, since it enhances the upside potential of risky loans while the downside risk is protected by limited liability.

Figure 5 illustrates the economic consequences of this volatility shock. As before, the dashed lines show what would happen if γ_t were to be held constant. The shock is big enough to entice banks to switch to risky loans, some of which will fail, increasing taxes and destroying bank equity. The story that follows is by now familiar. Consumption and savings/investment fall. Eventually, the shock dissipates and the falling capital stock raises R^s enough to make safe loans attractive again. As the solid lines illustrate, the Ramsey Planner would increase capital requirements just enough to eliminate the excessive risk taking. Under the Ramsey policy, there is no change in the expected return on safe equity or on S_t; the shock has absolutely no effect outside of financial markets, and only the capital requirement moves inside financial markets.

Takeaways: With no change in capital requirements, the effect of this shock on consumption and output is rather small; however, the shock itself was not large. Note that the path followed by the debt-to-GDP ratio and the investment price, Q_t, in the inefficient solution are not good indicators for the direction of optimal policy.

6.6 Sensitivity Analysis

The size of the optimal adjustments in capital requirements is strongly influenced by two parameters: the variance of the idiosyncratic shocks, τ, and the average penalty for taking a flier on a risky loan, ξ. In Figure 6, we focus on the TFP shock. The circles in these diagrams represent the baseline calibrations. The maximum adjustment in the optimal capital requirements is especially sensitive to increases in τ. At the outer range of the values of τ that we consider, we can boost the change in capital requirements to a more substantive 0.75 percent in response to a TFP shock that, at its peak, still reduces output by 1.5 percent, just as in Figure 3.

7 Implementable Buffer Rules

The Ramsey policy derived in Section 6 was in response to three different shocks, each of which was considered in isolation. In practice policymakers face a much more difficult challenge: the economy is actually driven by a multiplicity of shocks, all occurring at the same time; policymakers have to respond to the full stochastic structure of the economy. In our model, we can derive the Ramsey policy when the economy is hit by a full constellation of shocks, but it is implausible to think that policymakers would be able to implement it. So, in this section, we consider simple policy rules in which the capital requirement responds to one or two observable endogenous variables, and we ask which, if any, of these rules can closely mimic the actual Ramsey policy. Of particular interest will be Basel III's capital buffer rule in which capital requirements respond positively to the credit to GDP ratio.

This exercise is neither easy nor straightforward. The first step is to decide which shocks drive the macroeconomy, and we will see that this decision is not innocent: a different choice of shocks can alter results dramatically. Here, we will consider two calibrations which fit the U.S. data rather well. In Calibration 1, we use the two macroeconomic shocks – TFP and ISP (investment specific) – that were considered in the last section; in Calibration 2, we expand the set of shocks to include the volatility shock. The moments we match are the variances, correlation, and auto-covariances of chained real GDP, chained real private investment, and the implicit price deflator for chained investment (divided by price deflator for consumption).

The next step is to calibrate the shocks to make model moments match moments in the

U.S. data. We allow each shock to follow an auto-regressive process of order 1, and we need to size the persistence parameters and the standard deviations of the innovations. We also want to size the investment adjustment cost parameter, ϕ, and the habits parameter, κ. To do this, we use a SMM (simulated method of moments) procedure. For these calibrations, we are focusing on variances, covariances, and auto-covariances of all the observed variables, with the estimation sample starting in 1980. We experiment with the SMM optimal weighting matrix, and we match observed moments from bandpass-filtered data (selecting standard business cycle frequencies) against analogous moments simulated from a sample of 2000 model observations (also bandpass filtered).

Finally, it should be noted that we are also calculating and imposing the Ramsey policy for capital requirements in our model simulations. So, the model output gives us data for the optimal dynamic capital requirements, and model data are generated under the assumption that the optimal capital requirements are in place. With that assumption, there is no discernible difference in the targeted moments when we add the volatility shock under our Calibration 2. The Ramsey policy varies capital requirements to avoid excessive risk-taking episodes, otherwise leaving little imprint on the macroeconomy. As we shall see, this additional shock will lead to a substantial deterioration of the performance of simple rules. For illustration, we choose the same parameterization as for the earlier example in Figure 5.

7.1 Matching Moments, Shock Processes and Variance Decompostions

Tables 4 and 7 show that both calibrations are very good; model moments are close to data moments. Moreover, the values of the distance functions reported at the bottom of the tables show differences that are trivial, on the order of 2×10^{-7}.

Tables 2 and 3 show the calibrated shock processes and the variance decompositions associated with Calibration 1; Tables 5 and 6 report the analogous results for calibration 2. It may be interesting to note that the persistence parameter for the TFP shock in both calibrations is 0.79, which is somewhat lower than what is normally assumed in the RBC literature.[17] Finally, the parameters for consumption habits and investment adjustment costs that minimize the distance function are 0.93 and 0.06 respectively.

In Calibration 1, both shocks are persistent. But in the variance decompositions, the TFP shock does all of the work for GDP and investment; the ISP shock only matters for the investment price. Note also that the ISP shock explains all the variation in the Ramsey policy

[17]In most of the RBC literature, the persistence parameter is estimated by a simple auto-regression on TFP data.

setting, γ. In Calibration 2, all of the shocks are persistent. In the variance decompositions the TFP shock once again explains all of the variations in GDP and investment, while now the volatility shock explains the variation in the Ramsey policy settings.

7.2 Implementable Capital Buffer Rules

The Ramsey policy requires full knowledge of all the shocks, making its implementation virtually impossible in practice. Here, we focus on simple rules that may be able to mimic the optimal policy; these rules are based on one or two observable variables, and they are clearly implementable. The Basel III cyclical buffer, which runs off of the credit-to-GDP ratio, will be of particular interest. We will also compare these simple rules to more complex rules that are probably not implementable.

To derive the policy rules, we use data generated by our simulations. That is, we regress the Ramsey policy settings on one or more of the endogenous variables (and a constant). Then, we use a variety of measures to rank the alternative rules. The first, and perhaps the most obvious, measure is the R-square of the regression; the higher the R-square, the more closely the rule tracks the Ramsey settings. But there are other measures – performance measures – that focus on what the rule actually achieves. A good rule should minimize the frequency of excessive risk-taking episodes; the Ramsey policy eliminates them altogether. But recall that there is a tradeoff here. The frequency of episodes can also be minimized, or even eliminated, by simply setting the static capital requirement at a very high level. This cannot be the only performance measure that we consider since a very high capital requirement forces banks to limit the deposits they issue, and deposits are valued for their transactions services. So, the second performance measure is the average level of deposits that it achieves – the higher, the better.

Simple Rules Under Calibration 1

Table 8 reports our results for various policy rules under Calibration 1. The first column lists the variables in the rule; the second column gives the R-square for the rule's regression; the third and fourth columns show the regression coefficients; the fifth and sixth columns report the rule's performance measures: the average number of risk-taking quarters per 100 years and the average level of deposits when the static capital buffer is 10 basis points (that is, when the steady-state capital requirement is raised from 10 percent to 10.1 percent); and finally, the seventh and eighth columns report the performance measures when the static capital buffer is 30 basis points (or the steady-state capital requirement is raised to 10.3 percent).

The Ramsey Policy allows no risk-taking episodes, and the average level of deposits is 16.25. These performance measures – 0 and 16.25 – are the gold standard, the standard to which the implementable rules can only hope to aspire.

The best implementable rule for Calibration 1 has capital requirements responding to the investment price. The R-square is 0.96, so it tracks the Ramsey policy quite well. And this simple rule comes close to meeting the Ramsey performance standards – no risk-taking episodes, and an average level of deposits of 16.23 (with a static buffer of just 10 basis points). It is easy to see why this rule does so well. Figures 3 and 4 show that for both of the shocks that drive the economy, the investment price falls while the Ramsey capital requirement rises. Moreover, in Table 3, the ISP shock explains all the variation in the Ramsey requirement, and 92 percent of the variation in the investment price. So the investment price is a very good signal for what should be done with the capital requirement.

By contrast, the Basel rule does very poorly. The Basel III prescription is to tighten capital requirements when the credit-to-GDP ratio is rising and relax them when the ratio is falling. In Table 8, the R-square for this rule is only 0.25. Moreover, the number of risk-taking quarters per 100 years is very high when the steady-state capital requirement is 10.1 percent, and the average level of deposits is very low. Note also that the sign of the regression coefficient is wrong, at least from the perspective of the Basel III recommendations. In the next row, we impose a positive coefficient, and the results are even worse, as might have been expected.

Raising the steady-state capital requirement to 10.3% brings a huge improvement in the Basel rule. But, the higher steady-state capital requirement is doing all of the work here: the number of risk-taking quarters falls dramatically, and the level of deposits rises dramatically. The latter result may seem counter intuitive, since higher capital requirements force banks to decrease the proportion of loans that are funded by bank deposits. The answer to this puzzle is that the level of output and loans is lower during risk-taking episodes. Limiting the number of risk-taking episodes increases the average amount of credit that is extended, and this can raise the level of deposits even when deposits account for a lower fraction of the bank's funding.

So why does the Basel rule itself do so badly? Figures 3 and 4 show that for both of the shocks that drive the economy, the credit-to-GDP path reverses direction midway through, while the paths of the Ramsey capital requirement are monotonic. And from the variance decompositions reported in Table 3, the ISP shock drives the Ramsey capital requirements, while it only explains 41 percent of the variation in the credit-to-GDP ratio.

Some (incorrectly) interpret the Basel rule as saying that capital requirements should move pro-cyclically – increasing in booms and decreasing in recessions. But Table 8 shows that the GDP rule fares no better than the actual Basel rule. The R-square is virtually zero; so it is not tracking the Ramsey policy. And the performance measures are also bad.

The remaining rules are probably not implementable because of their informational requirements. The simplest is a rule that responds to the expected spread between the safe return and the deposit rate. This rule sounds sensible, given the discussion in Section 6.1, and indeed it has an R–square of 0.83; it tracks the Ramsey policy fairly well. However, its performance is so poor that risk-taking episodes can last beyond what our solution methods can accommodate, leading to convergence problems

The last two rules implausibly assume that the policymaker can observe the shocks and their innovations. Armed with all this information, the R-squares are 1.0. However, neither of these rules do any better than the simple investment price rule on the performance measures.

Simple Rules Under Calibration 2

Calibration 2 adds the volatility shock. Table 9 shows that the addition changes our results dramatically. Neither the investment price rule nor the Basel rule works well; they have low R-squares and poor performance measures unless the steady-state reserve requirement is raised from 10 percent to 11 percent. Here again, the work is being done by the static capital buffer, and not the rules themselves. The reason for the poor performance of these rules can be seen in the variance decompositions of Table 6. The volatility shock explains 97 percent of the variation in the capital requirement, but 0 percent of variation in the investment price and only 1 percent of the variation in the capital-to-GDP ratio.

The rule based upon the spread between the expected return on safe loans and the deposit rate does a better job of tracking the Ramsey policy, but once again the steady state is plagued by convergence problems. Only rules that assume an implausible amount of information, including the shocks processes and their innovations, come close to matching the performance of the Ramsey policy.

The Efficiency of Static Capital Buffers

The results reported in the previous sections seem to indicate that the steady-state capital requirement is an important instrument in the regulator's tool kit. Table 10 bears that out. Here, there are no rules, just static capital buffers. The last row gives the performance measures achieved by the Ramsey Planner. The first row with numbers reports the performance measures if the static capital requirement is raised from the 10 percent benchmark to 10.1 percent; they are not good. However, if the requirement is raised to 10.4 percent for Calibration 1, or 11.5 percent for Calibration 2, the results are almost as good as those achieved by the Ramsey Planner. This suggests that the regulator need not bother with dynamic capital requirements. If the static capital requirement is raised to 11.5 percent, the performance measures for both calibrations are very close to the optimal ones. Note, however, that the optimal level of the buffers depends on the calibration we are using.

Takeaways: Calibrations 1 and 2 show that changing the shock structure that drives the economy can radically alter the ability of simple rules to perform well. Simple rules, like the Basel rule, do not perform well for either calibration. However, eschewing policy rules and increasing the static capital requirement by as little as 1 percent nearly achieves the performance standards set by the Ramsey policy.

8 Conclusion

In our model, bank risk taking is endogenous, and the temptation to take excessive (or socially inefficient) risk is enabled by limited liability and government deposit insurance, which protect banks and depositors from the more extreme losses. Both macroeconomic shocks and market volatility shocks can trigger bouts of excessive risk taking by lowering the expected return on safer investments. Capital requirements can eliminate that temptation by making banks keep more skin in the game, but this may come at the cost of limiting liquidity-producing deposits.

We provide examples in which a Ramsey Planner would raise capital requirements in response to either cyclical booms or busts (depending upon the underlying shocks), and raise capital requirements in response to an increase in market volatility that has little consequence for the business cycle.

In practice, the policymaker's problem is more difficult than responding to a single well-identified shock. The policymaker has to respond to the full constellation of shocks that drive the economy. Accordingly, the informational requirements for a regulator are daunting, even in our stylized model where we only have two projects that banks can finance. In practice regulators would have to keep track of expected relative returns for a myriad possible projects.

We find it implausible to think that a policymaker could implement the optimal Ramsey policy in practice. In this environment, it is tempting to look for market indicators that might point the way to appropriate changes in the capital requirement. However, we showed that popular candidates – such as growth in the credit-to-GDP ratio – were unlikely to be a reliable indicators. To this end, we employed an SMM procedure to: (1) calibrate the shock processes that drive our model economy, (2) calculate the Ramsey policy in that environment, and (3) evaluate implementable policy rules against the Ramsey benchmark. Most policy rules fell into the risk-taking trap with an unfortunate frequency. Fortunately, we found that a small static buffer – slightly higher than the optimal steady-state capital requirement – avoided the Wile E. Coyote moments and achieved levels of deposits close to the Ramsey policy. Some finely tuned policy rules – such as the Basel III prescriptions – sound like they make sense, but they do more harm than good in our model. Fine tuning capital requirements seems exceedingly risky; the Hippocratic Oath – First, do no harm – may be an appropriate guide for well-intentioned regulators.

Table 1: Parameters

	Value	Description	
Conventional			
β	0.99	Discount rate	
α	0.3	Capital share in production	
ϱ_c	1.1	Elasticity of substitution for consumption	
δ	0.025	Depreciation rate	
ς_d	1.1	Interest rate elasticity of supply of deposits	
Specific			*Target/Explanation*
τ	0.05521	Standard deviation of idiosyncratic shock	$\frac{\text{Debt}}{\text{EBITDA}} = 6$
ξ	0.00076	Minus mean of idiosyncratic shock	Cap. requirement= 10%
ς_0	0.015	Relative weight on liquidity in the utility function	Quarterly rate on bank debt= 0.86%
f	0.0055	Linear Cost of Banking	$R^s - R^d = 2.26\%$
ϕ	0.06	Investment adjustment costs	estimated by SMM
κ	0.93	Habits	estimated by SMM
$\underline{\sigma}$	0.01	Minimum risk that banks can take	needed for numerical solution method
$\bar{\sigma}$	0.99	Maximum risk that banks can take	needed for numerical solution method

Table 2: Calibration 1, Shock Processes

	AR(1) param.	Innov. St. Dev.
TFP	0.79	0.0093
ISP	0.95	0.0052
Distance Function	0.0012289861	

Table 3: Calibration 1, Variance Decomposition

	var(GDP)	var(invest.)	var(invest. p.)	var(gamma)	var(credit/GDP)
TFP	100	99	8	0	59
ISP	0	1	92	100	41

Table 4: Calibration 1, Matching Moments

	Data	Model
Var(GDP)	0.92	0.97
Corr(GDP,Investment)	0.96	1.00
Corr(GDP,Investment Price)	0.08	0.08
Var(Investment)	27.68	27.68
Corr(Investment,Investment Price)	0.02	0.06
Var(Investment Price)	0.40	0.38
Autocorr(GDP)	0.93	0.88
Autocorr(Investment)	0.93	0.88
Autocorr(Investment Price)	0.87	0.88

Table 5: Calibration 2: Shock Processes

	AR(1) param.	Innov. St. Dev.
TFP	0.79	0.0093
ISP	0.95	0.0052
Volatility	0.80	0.0015
Distance Function	0.0012289856	

Table 6: Calibration 2, Variance Decomposition

	var(GDP)	var(invest.)	var(invest. p.)	var(gamma)	var(credit/GDP)
TFP	100	100	8	0	65
ISP	0	0	92	2	35
Volatility	0	0	0	98	0

Table 7: Calibration 2, Matching Moments

	Data	Model
Var(GDP)	0.92	0.97
Corr(GDP,Investment)	0.96	1.00
Corr(GDP,Investment Price)	0.08	0.08
Var(Investment)	27.68	27.68
Corr(Investment,Investment Price)	0.02	0.06
Var(Investment Price)	0.40	0.38
Autocorr(GDP)	0.93	0.88
Autocorr(Investment)	0.93	0.88
Autocorr(Investment Price)	0.87	0.88

Table 8: Simple Rules with Calibration 1

		Regression coefficients		**Static buffer = 10 basis points**		**Static buffer = 30 basis points**	
Simple rule	**R square**	**First variable**	**Second variable**	**Quarters with excessive risk-taking (per 100 years)**	**Average deposit under simple rule**	**Quarters with excessive risk-taking (per 100 years).**	**Average deposit under simple rule.**
Invest. p. (best state variable)	0.960	-0.087		0	16.23	0	16.20
Expected banking spread	0.881	0.842		115.6	11.50	0	16.20
GDP	0.002	-0.001		149.6	10.21	10.4	15.79
Credit/GDP	0.250	-0.005		149.2	10.18	4.4	16.02
Credit/GDP wih positive coef		0.005		158.8	9.87	38	14.68
Expected safe return and deposit rate	0.826	594.284	-594.312	convergence problems	convergence problems	convergence problems	convergence problems
All shock processes, innovations, expected safe return and deposit rate	1.000	Too many to show		0	16.23	0	16.20
All shock processes, innovations, and lagged capital requirement	1.000	Too many to show		0	16.23	0	16.20

Table 9: Simple Rules with Calibration 2

Simple Rule	R Square	First variable	Second variable	Quarters with excessive risk-taking (per 100 years)	Average deposit under simple rule	Quarters with excessive risk-taking (per 100 years).	Average deposit under simple rule.	Number quarters with excessive risk-taking (per 100 years)..	Average deposit under simple rule..
	Regression coeffiecients			**Static buffer = 10 basis points**		**Static buffer = 50 basis points**		**Static buffer = 100 basis points**	
1. Invest. p.	0.043	-0.066		195.6	8.273	69.6	13.297	6.0	15.830
2. Expected banking spread	0.613	0.773		211.2	7.647	77.6	12.991	6.8	15.802
3. GDP	0.000	-0.001		210.8	7.697	79.6	12.903	6.8	15.805
4. Credit/GDP	0.016	-0.005		208.4	7.777	76.8	13.027	7.2	15.788
5. Credit/GDP wih positive coef		0.005		Convergence problems		83.2	12.780	6.8	15.805
6. Expected safe return and deposit rate	0.974	861.783		Convergence problems		Convergence problems		Convergence problems	
7. All shock processes, innovations, expected safe return and deposit rate	1.000	Too many to show	-861.892	0	16.223	0.0	16.151	0	16.061
8. All shock processes, innovations, and lagged capital requirement	1.000	Too many to show		0	16.223	0.0	16.151	0	16.061

Table 10: The Efficiency of Static Buffers

Static Buffer	**Calibration 1 (excludes volatility shocks)** Number of quarters with excessive risk-taking (per 100 years)	**Calibration 1 (excludes volatility shocks)** Average deposit	**Calibration 2 (includes volatility shocks)** Number of quarters with excessive risk-taking (per 100 years)	**Calibration 2 (includes volatility shocks)** Average deposit
10 bp	149.2	10.261	210.8	7.678
20 bp	66.8	13.526	172.0	9.216
30 bp	10.8	15.785	140.8	10.479
40 bp	0	16.189	108.8	11.784
50 bp	0	16.171	79.2	12.920
100 bp	0	16.081	6.8	15.805
150 bp	0	15.991	0	15.991
Optimal Rule	0	16.251	0	16.241

Figure 1: **Higher Capital Requirement Shock**

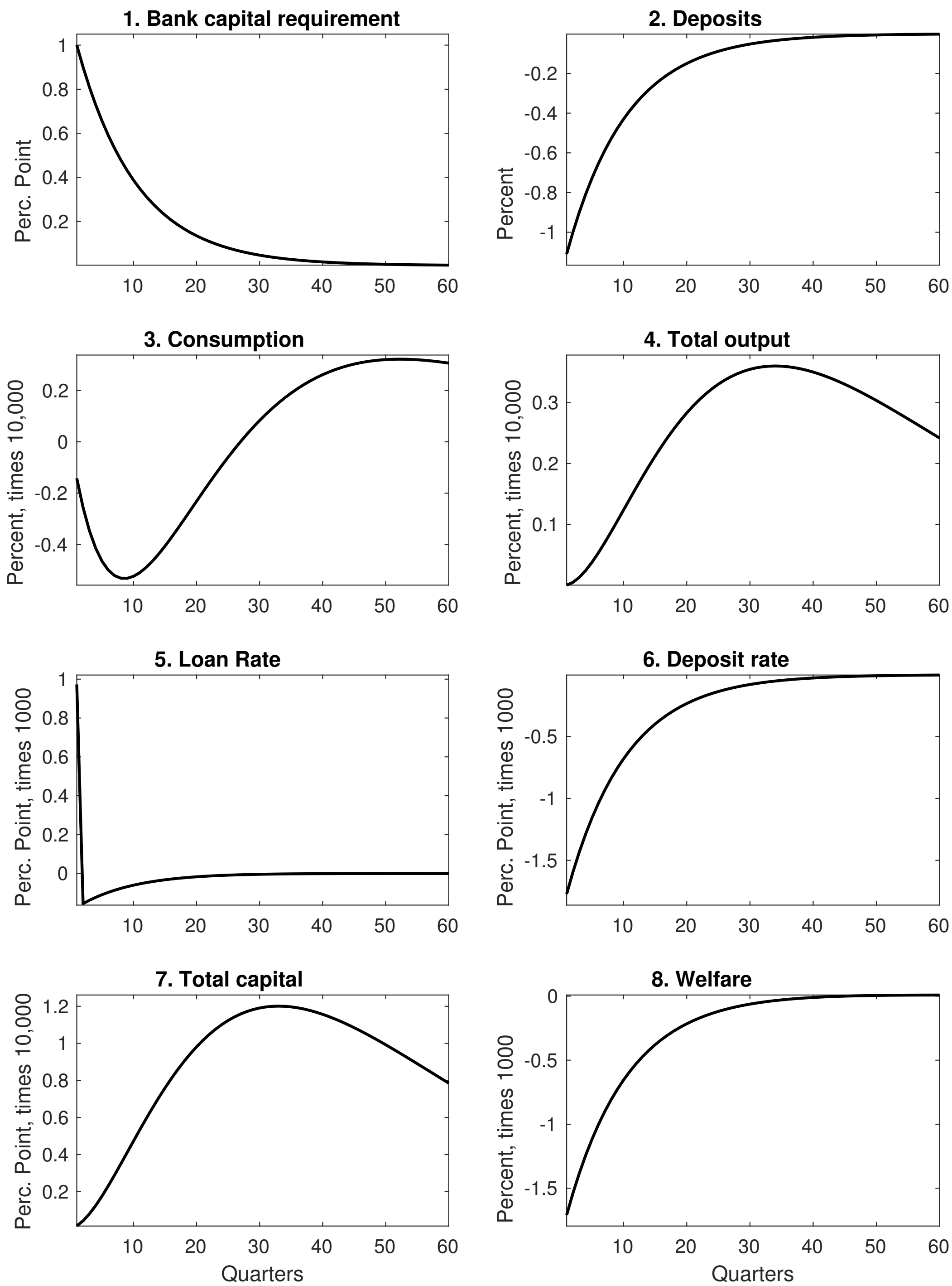


Figure 2: **Capital Requirement Shocks**

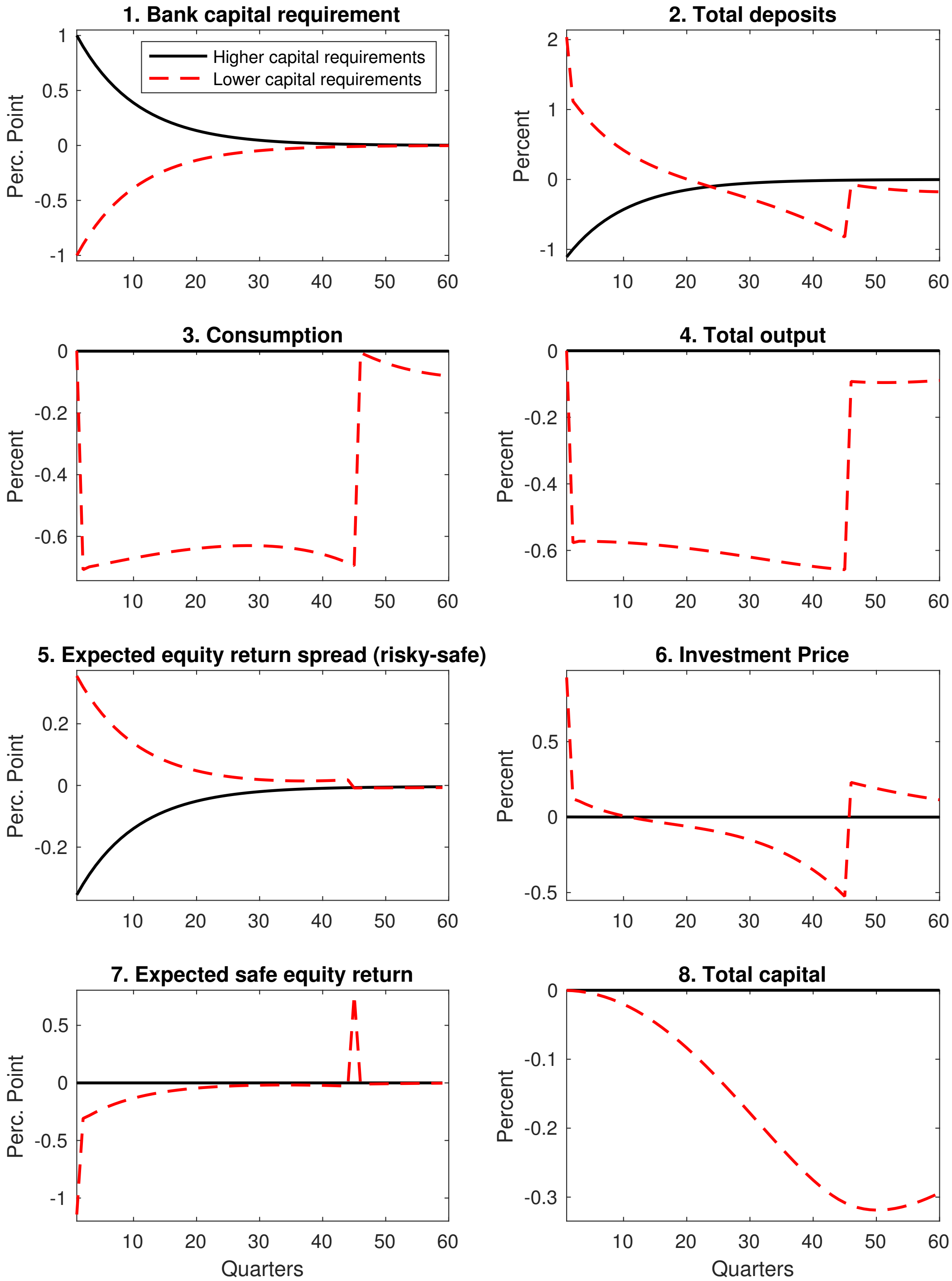


Figure 3: **Negative TFP Shock**

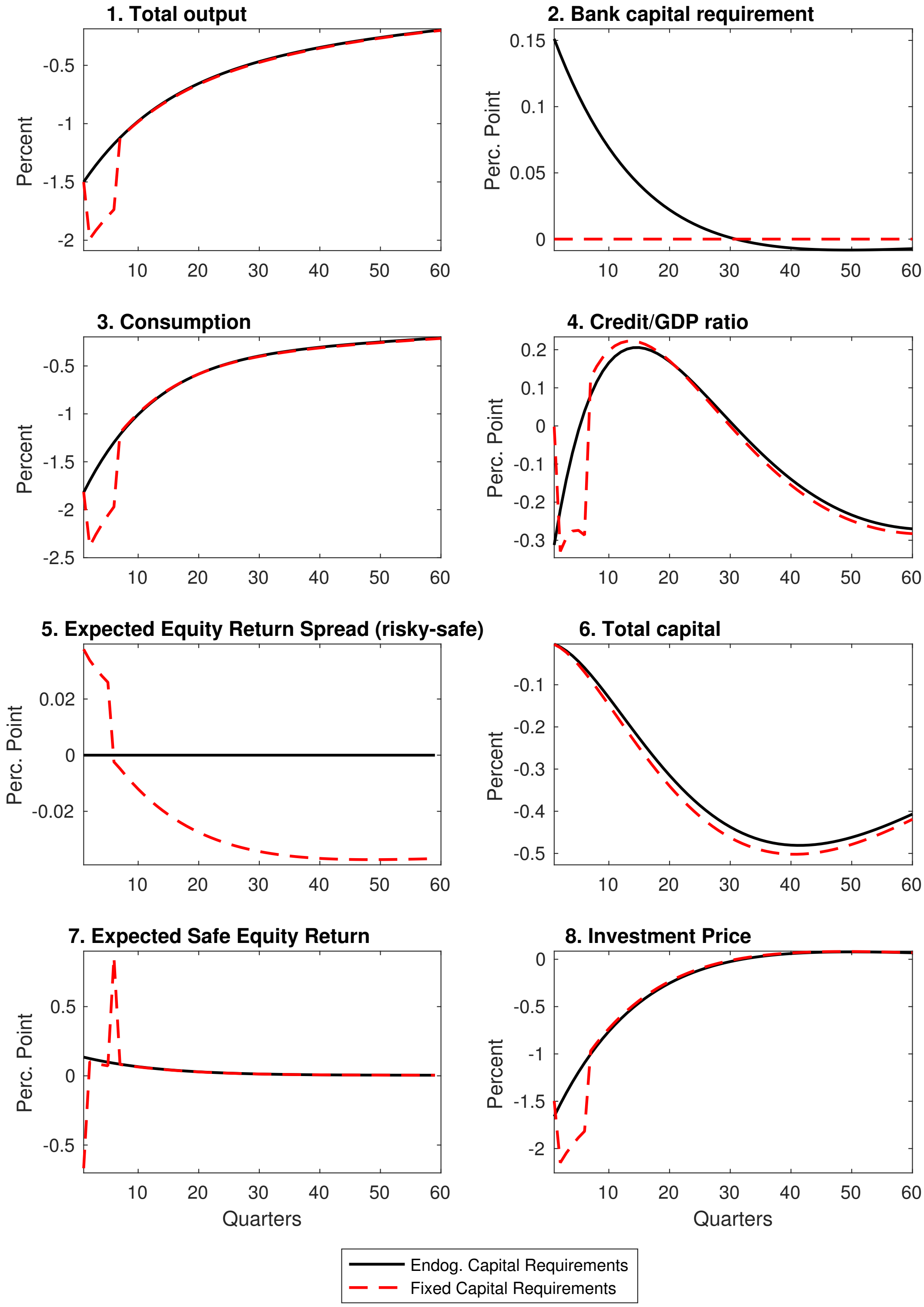


Figure 4: **Positive Investment Shock**

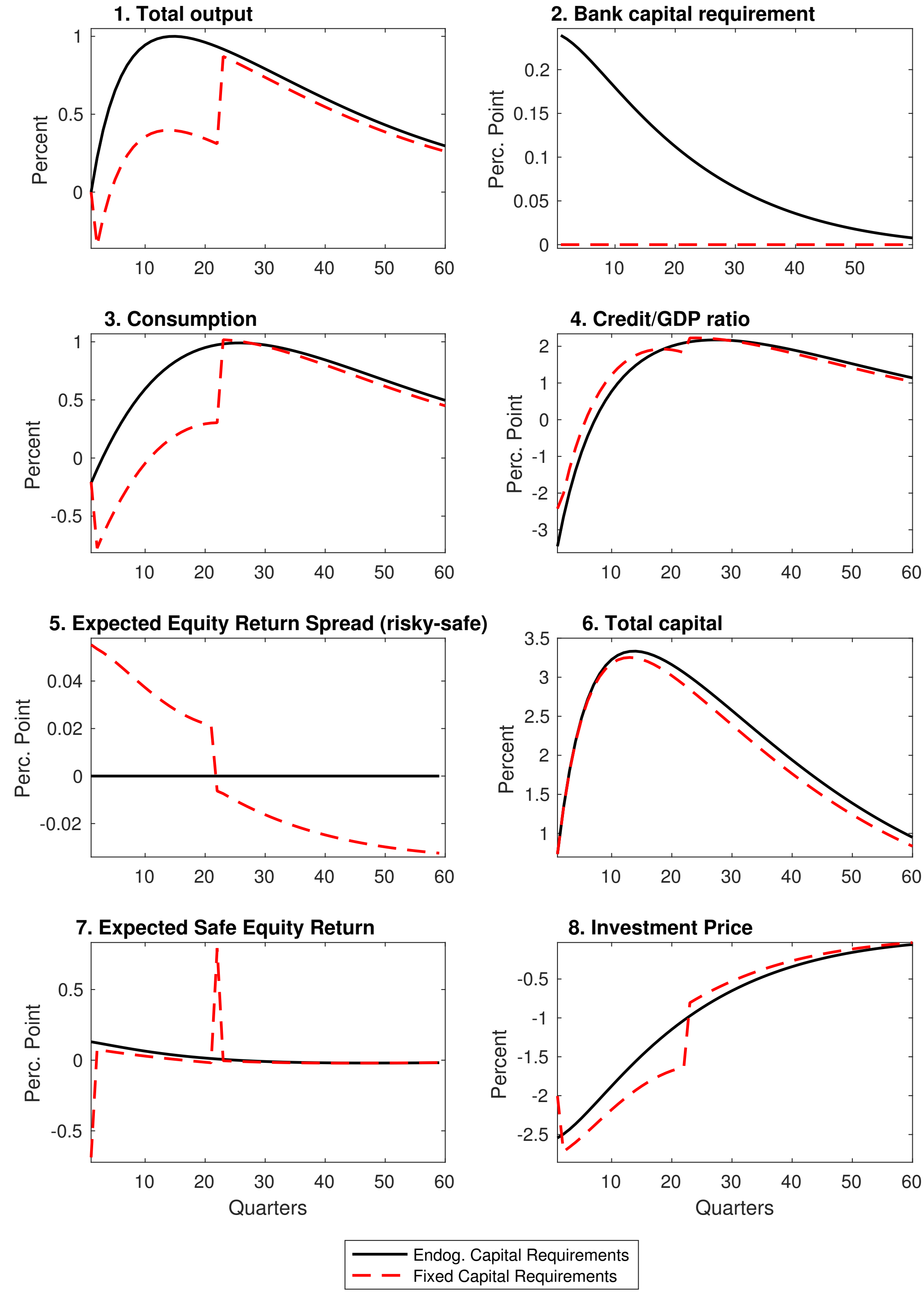


Figure 5: **Volatility Shock**

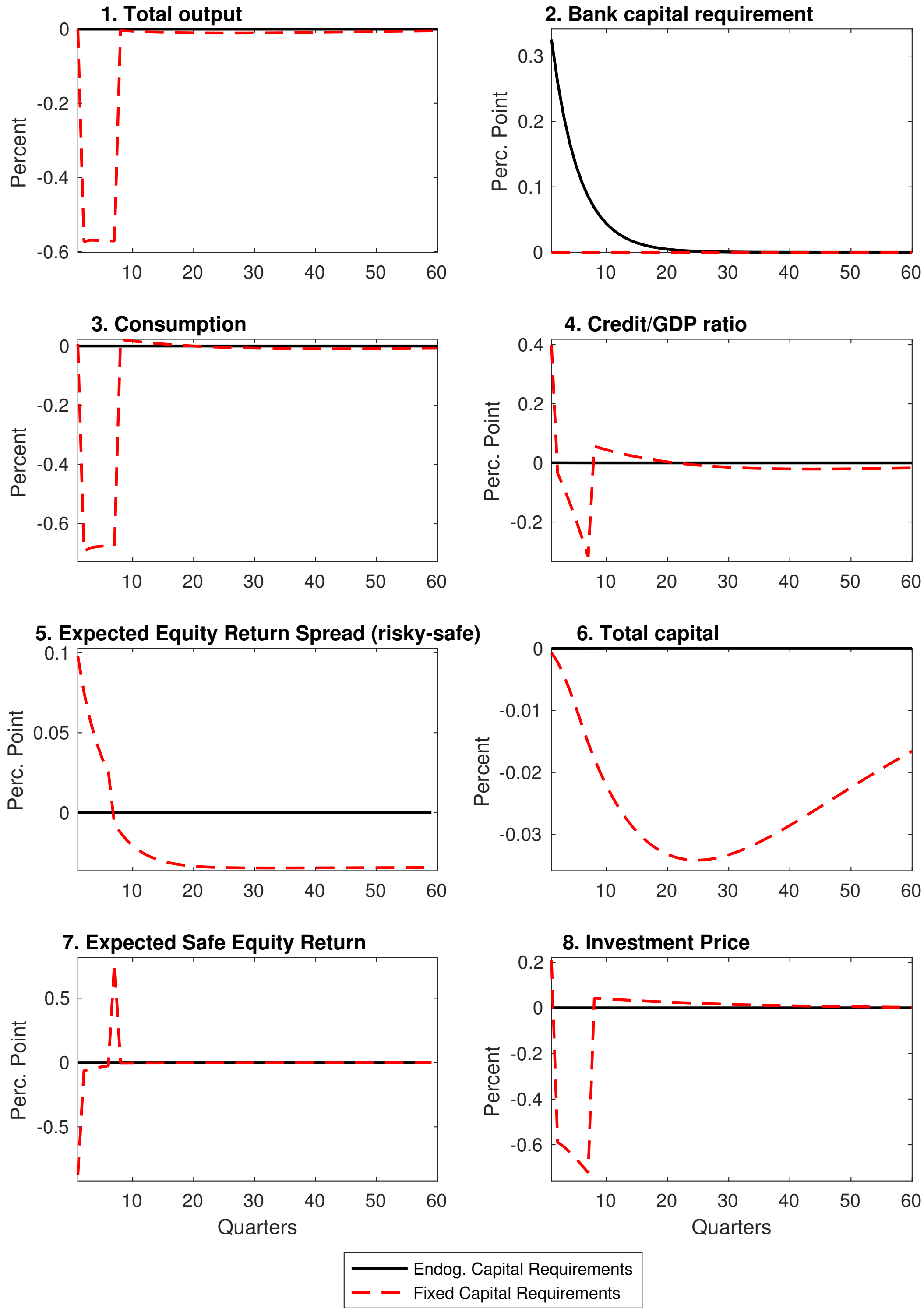
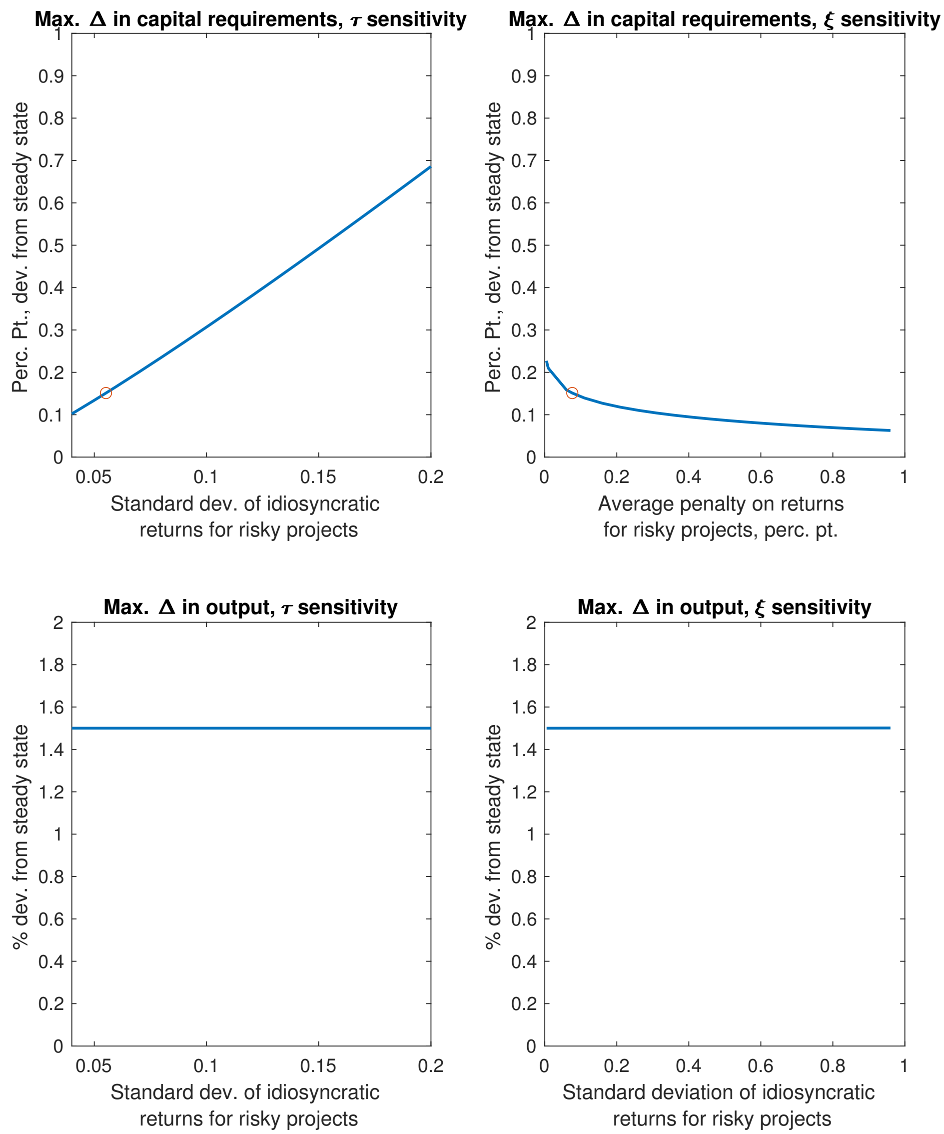


Figure 6: **Sensitivity Analysis, TFP Shock**



9 References:

Altig, D., L. Christiano, M., Eichenbaum, J. Linde (2011). Firm-specific capital, nominal rigidities and the business cycle. *Review of Economic Dynamics* 14 (2), 225–247.

Angeloni, I and E. Faia (2013). Capital Regulation and Monetary Policy with Fragile Banks. *Journal of Monetary Economics* 60 (3), 311–324.

Begenau, J. (2019). Capital Requirements, Risk Choice, and Liquidity Provision in a Business Cycle Model. *Journal of Financial Economics*, forthcoming.

Christiano, L. J., Motto, R., and Rostagno, M. (2014). Risk shocks. American Economic Review, 104(1), 27-65.

Collard, F., H. Dellas, B. Diba, and O. Loisel (2017). Optimal Monetary and Prudential Policies. *American Economic Journal: Macroeconomics* 9 (1), 40–87.

Davydiuk, T. (2017). Dynamic Bank Capital Requirements. 2017, mimeo.

Elenev, V., T. Landvoigt, and S. V. Nieuwerburgh (2018). A Macroeconomic Model with Financially Constrained Producers and Intermediaries. NBER Working Papers 24757.

Faria-e-Castro, M. (2019). A Quantitative Analysis of Countercyclical Capital Buffers. FRB St. Louis Working Paper, (2019-8).

Gersbach, H. and J. C. Rochet. (2017). Capital Regulation and Credit Fluctuations. *Journal of Monetary Economics* 90, 113–124.

Gertler, M. and N. Kiyotaki (2015). Banking, Liquidity, and Bank Runs in an Infinite Horizon Economy. American Economic Review 105 (7), 2011-2043.

Gertler, M., Kiyotaki, N., & Prestipino, A. (2020). A macroeconomic model with financial panics. The Review of Economic Studies, 87(1), 240-288.

Gomes, J., M. Grotteria, and J. Wachter (2018). Foreseen Risks. NBER Working Paper No. 25277.

Guerrieri, L. and M. Iacoviello (2015). OccBin: A toolkit for solving dynamic models with occasionally binding constraints easily. *Journal of Monetary Economics* 70 (C), 22–38.

Jordà, O., M. Schularick, and A. M., Taylor (2016). Macro Financial History and the New Business Cycle Facts. *NBER Macroeconomics Annual* 31, University of Chicago Press, 213-263.

Kashyap, A. K. and J. C. Stein (2004). Cyclical Implications of the Basel II Capital Standards. *Economic Perspectives - Federal Reserve Bank of Chicago* 28 (1), 18–33.

Malherbe, F. (2020). Optimal Capital Requirements Over the Business and Financial Cycles. *American Economic Journal: Macroeconomics*, forthcoming.

Martinez-Miera, D. and J. Suarez (2014). A Macroeconomic Model of Endogenous Systemic Risk Taking, mimeo.

Mendicino, C., K. Nikolov, J. Suarez, and D. Supera (2018). Optimal Dynamic Capital Requirements. Journal of Money, Credit and Banking, 50, 1271–1297.

Mian, A., A. Sufi, and E. Verner (2017). Household Debt and Business Cycles Worldwide. *Quarterly Journal of Economics* 132, 1755-1817.

Repullo, R. and J. Suarez (2013). The Procyclical Effects of Bank Capital Regulation. *Review of Financial Studies* 26 (2), 452–490.

Schularick, M., and A. M. Taylor (2012). Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008. *American Economic Review* 102, 1029-1061.

Van den Heuvel, S. J. (2008). The Welfare Cost of Bank Capital Requirements. *Journal of Monetary Economics* 55 (2), 298–320.

Van den Heuvel, S. J. (2019). The Welfare Effects of Bank Liquidity and Capital Requirements, 2019 Meeting Papers 325, Society for Economic Dynamics.

For Online Publication

Online Appendix for

"A Static Capital Buffer is Hard To Beat"

Matthew Canzoneri Behzad Diba

Luca Guerrieri Arsenii Mishin

November 2020

A The Bank's Problem

A.1 Baseline: First-Order Conditions

Substituting $d_t = l_t - e_t$ into equation (20) and writing $\mathrm{d}G(\varepsilon_{t+1})$ explicitly turn the objective into:

$$\max_{l_t,e_t,\sigma_t} E_t \left\{ \psi_{t,t+1} \left[\int_{\varepsilon^*_{t+1}}^{\infty} \left(\left(R^s_{t+1} + \sigma_t \frac{\varepsilon_{t+1}}{Q_t} \right) l_t - R^d_t \left(l_t - e_t \right) \right) \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \, \mathrm{d}\varepsilon_{t+1} \right] - e_t \right\},$$

subject to

$$e_t \geq \gamma_t l_t,$$
$$l_t \geq 0,$$
$$\underline{\sigma} \leq \sigma_t \leq \bar{\sigma},$$

where $\psi_{t,t+1} = \beta \frac{\lambda_{ct+1}}{\lambda_{ct}}$ is the stochastic discount factor and $\varepsilon^*_{t+1} = \left(\frac{R^d_t - R^s_{t+1}}{\sigma_t} - \frac{R^d_t e_t}{\sigma_t l_t} \right) Q_t$ is the shield of limited liability. Note that we expressed ε^*_{t+1} from $\left(R^s_{t+1} + \sigma_t \frac{\varepsilon^*_{t+1}}{Q_t} \right) l_t - R^d_t \left(l_t - e_t \right) = 0$ to get the lower limit of the integral.

Append the Lagrangian 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:

$$\frac{\partial \mathcal{L}}{\partial l_t} = E_t \left[\psi_{t,t+1} \overbrace{\left(\left(R^s_{t+1} + \sigma_t \left(\frac{R^d_t \left(l_t - e_t \right)}{\sigma_t l_t} - \frac{R^s_{t+1}}{\sigma_t} \right) \right) l_t - R^d_t \left(l_t - e_t \right) \right)}^{=0} \cdot \frac{\partial \varepsilon^*_{t+1}}{\partial l_t} \right] + \chi_{2t} +$$

$$E_t \left[\int_{\varepsilon^*_{t+1}}^{\infty} \psi_{t,t+1} \frac{\partial}{\partial l_t} \left(\left(R^s_{t+1} + \sigma_t \frac{\varepsilon_{t+1}}{Q_t} \right) l_t - R^d_t \left(l_t - e_t \right) \right) \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \, \mathrm{d}\varepsilon_{t+1} \right] - \gamma \chi_{1t} = 0,$$

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

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

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

We are using the Leibniz integral rule above to find the partial derivatives of the profit function. 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.

Next, express the integrals in the first-order conditions above using the erf function, wherever possible. Note that in order to make the next expressions more neat we omit the stochastic discount factor and the expectation operator from consideration. We include them in the final exposition.

Work on $\frac{\partial}{\partial l_t}$:

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

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

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

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

Break the calculation of the integral into two parts.

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

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 $\int_a^b f(\varphi(t))\varphi'(t)dt$. 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 $\sqrt{2}\tau$.

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

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

$$-\frac{\sqrt{2}\tau}{2\sqrt{\pi}} e^{-v^2}\Bigg|_{\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}^{\infty} - \frac{\xi}{\sqrt{\pi}} \left[\int\limits_0^{\infty} e^{-v^2}\,\mathrm{d}v - \int\limits_0^{\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}} e^{-v^2}\,\mathrm{d}v\right] =$$

$$0 + l_t \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)^2} -$$

$$\frac{\xi}{\sqrt{\pi}}\left[\frac{\sqrt{\pi}}{2}\mathrm{erf}(\infty) - \frac{\sqrt{\pi}}{2}\mathrm{erf}\left(\frac{\left(R_t^d\left(l_t-e_t\right)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right] =$$

$$\frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)^2} - \frac{\xi}{2}\left[1 - \mathrm{erf}\left(\frac{\left(R_t^d\left(l_t-e_t\right)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right],$$

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

Let's express $\int\limits_{\left(\frac{R_t^d-R_{t+1}^s}{\sigma_t}-\frac{R_t^d e_t}{\sigma_t l_t}\right)Q_t}^{\infty} \left(\frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\right) \mathrm{d}\varepsilon_{t+1}$ in terms of the error function.

Again, use the transformation $v = \frac{\varepsilon_{t+1}+\xi}{\sqrt{2}\tau}$ or $\varepsilon_{t+1} = v\sqrt{2}\tau - \xi$

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

$$\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{\left(R_t^d\left(l_t-e_t\right)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right).$$

Therefore,

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

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

$$E_t\left[\left(R_{t+1}^s-R_t^d\right)\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right)\right]=$$

$$E_t\left[\frac{\sigma_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)^2}+\right.$$

$$\left.\left(\frac{R_{t+1}^s-\frac{\sigma_t\xi}{Q_t}-R_t^d}{2}\right)\left[1-\operatorname{erf}\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right]\right].$$

Similarly, work on $\frac{\partial}{\partial e_t}$

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

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

In sum, the FOCs can be written as follows:

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

There are complementary slackness conditions which can be described by:

$$\begin{aligned}(e_t-\gamma l_t)\,\chi_{1t} &= 0,\\ l_t\chi_{2t} &= 0.\end{aligned}$$

A.2 Proof of Proposition 1

Equations (18) and (19) can be expressed as

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

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

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

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

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

where

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

$$\frac{\partial}{\partial D_t} = \varsigma_0 D_t^{-\varsigma_d} - \lambda_{ct} + E_t \beta \lambda_{ct+1} R_t^d = 0,$$

Use $E_t \beta \lambda_{ct+1} R_{t+1}^{e,i} + \zeta_t^i = \lambda_{ct}$ (that comes from the household's FOCs with respect to e_t^i for each $i \in \{s, r\}$) to substitute for λ_{ct} in equation (17) . We get:

$$E_t \left\{ \beta \lambda_{ct+1} \left[R_t^d - R_{t+1}^{e,i} \right] \right\} = -\varsigma_0 D_t^{-\varsigma_d} + \zeta_t^i.$$

Note that $\varsigma_0 D_t^{-\varsigma_d} > 0$ under the usual (and mild) assumptions on the preferences for liquidity. Moreover, the Lagrangian 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). The latest expression is transformed into the following after dividing it by λ_{ct}:

$$\underbrace{E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[R_t^d - R_{t+1}^{e,i} \right] \right\}}_{=\Psi_2^*} - \frac{\zeta_t^i}{\lambda_{ct}} = -\frac{\varsigma_0 D_t^{-\varsigma_d}}{\lambda_{ct}} < 0.$$

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

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

Hence, $\chi_{1t}^i > 0$. □

A.3 Combined First-Order Conditions

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

Since $\chi_{1t} > 0$, multiply the second equation by γ_t and add it to the first equation using $\frac{e_t}{l_t} = \gamma_t$. Therefore, the FOCs can be combined into:

$$E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[\frac{\sigma_t}{Q_t} \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^d(1-\gamma_t) - R_{t+1}^s\right)Q_t + \xi\sigma_t}{\sigma_t\sqrt{2}\tau} \right)^2} + \right. \right.$$

$$\left. \left. \frac{1}{2}\left(R_{t+1}^s - \frac{\sigma_t \xi}{Q_t} - R_t^d \right) \left[1 - \operatorname{erf}\left(\frac{\left(R_t^d(1-\gamma_t) - R_{t+1}^s\right)Q_t + \xi\sigma_t}{\sigma_t\sqrt{2}\tau} \right) \right] \right] \right\} = \gamma_t - \chi_{2t},$$

$$\chi_{2t} l_t = 0.$$

A.4 Zero-Profit Condition

Consider the zero-profit condition under all states of nature. Since there is no agency problem between banks and households, this condition captures the fact that all the profits (or losses) are distributed to equity holders after realization of shocks at the beginning of each period. In each aggregate state, banks whose investments in risky firms pan out will have returns that satisfy on average (over the realizations of the idiosyncratic shock) $\left[\left(R_{t+1}^s + \frac{\sigma_t}{Q_t} \right) l_t - R_t^d \left(l_t - e_t \right) \right] - \int R_{t+1,b}^e(b) \cdot e_t = 0$, where the bounds of the integral are chosen such that we integrate over banks for which the profit is non-negative, while banks whose risky investments earn low (negative) returns will have $R_{t+1,b}^e = 0$. Therefore,

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

$$\int_{-\infty}^{\left(\frac{R_t^d(1-\gamma_t) - R_{t+1}^s}{\sigma_t}\right)Q_t} 0 \cdot \frac{1}{\sqrt{2\pi\tau^2}} \mathrm{e}^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \, \mathrm{d}\varepsilon_{t+1} =$$

$$\frac{1}{e_t} \int_{\left(\frac{R_t^d(1-\gamma_t) - R_{t+1}^s}{\sigma_t}\right)Q_t}^{\infty} \left(R_{t+1}^s l_t - R_t^d d_t \right) \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \, \mathrm{d}\varepsilon_{t+1} +$$

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

$$\frac{1}{e_t}\left[\left(R^s_{t+1}l_t - R^d_t d_t\right)\frac{1}{2}\left(1-\text{erf}\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right)+\right.$$
$$\left.\frac{\sigma_t l_t}{Q_t}\left(\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)^2}-\frac{\xi}{2}\left[1-\text{erf}\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right]\right)\right]=$$

$$\frac{l_t}{e_t}\left\{\frac{\sigma_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)^2}+\right.$$
$$\left.\frac{1}{2}\left(R^s_{t+1}-\frac{\sigma_t\xi}{Q_t}-R^d_t\left(1-\gamma_t\right)\right)\left[1-\text{erf}\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right]\right\}.$$

Since $\frac{l_t}{e_t}=\frac{1}{\gamma_t}$, we can rewrite the latter condition as (using that it holds for each $i\in\{s,r\}$):

$$R^{e,i}_{t+1}=\frac{\frac{\sigma^i_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma^i_t}{\sigma^i_t\sqrt{2}\tau}\right)^2}+\frac{1}{2}\left(R^s_{t+1}-\frac{\sigma^i_t\xi}{Q_t}-R^d_t(1-\gamma_t)\right)\left[1-\text{erf}\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma^i_t}{\sigma^i_t\sqrt{2}\tau}\right)\right]}{\gamma_t}.$$

Note that the combined FOC from Appendix A.3 can be expressed as:

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\frac{\sigma^i_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma^i_t}{\sigma^i_t\sqrt{2}\tau}\right)^2}+\right.\right.$$
$$\left.\left.\frac{1}{2}\left(R^s_{t+1}-\frac{\sigma^i_t\xi}{Q_t}-R^d_t\right)\left[1-\text{erf}\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma^i_t}{\sigma^i_t\sqrt{2}\tau}\right)\right]\right]\right\}=$$
$$\gamma_t-\chi^i_{2t}=\gamma_t\left(E_t\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^{e,i}_{t+1}+\frac{\zeta^i_t}{\lambda_{ct}}\right)-\chi^i_{2t},$$

where we substitute for 1 from Household's FOC with respect to two types of equity: $\beta E_t\frac{\lambda_{ct+1}}{\lambda_{ct}}R^{e,i}_{t+1}=1-\frac{\zeta^i_t}{\lambda_{ct}}$.

Notice that $l^i_t>0$ implies both $\chi^i_{2t}=0$ and $\zeta^i_t=0$ which say that the zero-profit condition implies the FOC.

A.5 Expression of Expected Dividends

Expected dividends (valued on date t) are defined as

$$\Omega\left(\mu_t,\sigma_t;\ l_t,\ d_t,\ e_t\right)=$$

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

We have already calculated all the necessary integrals in Appendix A.1. Therefore,

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

A.6 Linear Cost of Banking: FOCs of Banks

We use $\left(R_{t+1}^s+\sigma_t\frac{\varepsilon_{t+1}^*}{Q_t}\right)l_t-R_t^d d_t-fl_t=0$ to get $\varepsilon_{t+1}^*=\left(\frac{fl_t+R_t^d(l_t-e_t)}{\sigma_t l_t}-\frac{R_{t+1}^l}{\sigma_t}\right)Q_t$. Conditional on the optimal choice of σ_t, the first-order conditions are:

$$E_t\left[\int\limits_{\varepsilon_{t+1}^*}^{\infty}\psi_{t,t+1}\frac{\partial}{\partial l_t}\left(\left(R_{t+1}^s+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}\right)l_t-R_t^d\left(l_t-e_t\right)-fl_t\right)\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}\right]+$$
$$\chi_{2t}-\gamma\chi_{1t}=0,$$

$$E_t\left[\int\limits_{\varepsilon_{t+1}^*}^{\infty}\psi_{t,t+1}\frac{\partial}{\partial e_t}\left(\left(R_{t+1}^s+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}\right)l_t-R_t^d\left(l_t-e_t\right)-fl_t\right)\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}\right]-$$
$$1+\chi_{1t}=0.$$

The derivations are similar to the ones described in Appendix A.1. The only difference is that the lower bound of the integral now contains the additional term fl_t. Hence, adding ξ to the lower limit of the integral and dividing the result by $\sqrt{2}\tau$ make the terms in the final expressions. Moreover, note that we should carry f in the expressions of the FOC with respect to l_t. In sum, the FOCs can be written as follows:

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\frac{\sigma_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(f+R_t^d\left(1-\frac{e_t}{l_t}\right)-R_{t+1}^s\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)^2}+\right.\right.$$

$$\left.\left.\left(\frac{R_{t+1}^s-\frac{\sigma_t\xi}{Q_t}-R_t^d-f}{2}\right)\left[1-\operatorname{erf}\left(\frac{\left(f+R_t^d\left(1-\frac{e_t}{l_t}\right)-R_{t+1}^s\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right]\right]\right\}+\chi_{2t}=\gamma\chi_{1t},$$

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

B The Non-Financial Firm's Problem

B.1 Safe firms

Let π_{t+1}^s denote the revenue of a safe firm in period $t+1$ net of expenses:

$$\pi_{t+1}^s=y_{t+1}^s+(1-\delta)Q_tk_{t+1}^s-W_{t+1}h_{t+1}^s-R_{t+1}^sl_t^{f,s}.$$

In this notation, the problem of the safe firm is to

$$\max_{l_t^{f,s},k_{t+1}^s}E_t\left\{\max_{h_{t+1}^s}\pi_{t+1}^s\right\}.$$

The first-order condition for $\max_{h_{t+1}^s}\pi_{t+1}^s$ is $\frac{\partial\pi_{t+1}^s}{\partial h_{t+1}^s}=0$. It implies that

$$W_{t+1}=\frac{\partial y_{t+1}^s}{\partial h_{t+1}^s}=(1-\alpha)\frac{y_{t+1}^s}{h_{t+1}^s}=(1-\alpha)A_{t+1}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^\alpha, \tag{B.1}$$

$$h_{t+1}^s=(1-\alpha)\frac{y_{t+1}^s}{W_{t+1}}=(1-\alpha)\frac{A_{t+1}\left(k_{t+1}^s\right)^\alpha\left(h_{t+1}^s\right)^{1-\alpha}}{W_{t+1}}. \tag{B.2}$$

Accordingly, the safe firm's Lagrangian is:

$$\mathcal{L}^{\text{safe}}=E_t\left\{A_{t+1}\left(k_{t+1}^s\right)^\alpha\left(h_{t+1}^s\right)^{1-\alpha}+(1-\delta)Q_{t+1}k_{t+1}^s-W_{t+1}h_{t+1}^s-R_{t+1}^sl_t^{f,s}\right\}+$$

$$\lambda_{ht}^sE_t\left\{(1-\alpha)\frac{A_{t+1}\left(k_{t+1}^s\right)^\alpha\left(h_{t+1}^s\right)^{1-\alpha}}{W_{t+1}}-h_{t+1}^s\right\}+\lambda_{lt}^s\left(l_t^{f,s}-Q_tk_{t+1}^s\right).$$

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

$$\frac{\partial \mathcal{L}^{\text{safe}}}{\partial l_t^{f,s}} = -E_t\left[R_{t+1}^s\right] + \lambda_{lt}^s = 0,$$

$$\frac{\partial \mathcal{L}^{\text{safe}}}{\partial k_{t+1}^s} = E_t\left[\alpha\frac{y_{t+1}^s}{k_{t+1}^s} + (1-\delta)Q_{t+1}\right] + \lambda_{ht}^s\left(1-\alpha\right)\alpha E_t\left[\frac{A_{t+1}}{W_{t+1}}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^{\alpha-1}\right] - \lambda_{lt}^s Q_t = 0,$$

$$\frac{\partial \mathcal{L}^{\text{safe}}}{\partial h_{t+1}^s} = (1-\alpha)\frac{A_{t+1}\left(k_{t+1}^s\right)^{\alpha}\left(h_{t+1}^s\right)^{1-\alpha}}{W_{t+1}} - W_{t+1} + \lambda_{ht}^s\left[(1-\alpha)^2\frac{A_{t+1}}{W_{t+1}}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^{\alpha} - 1\right] = 0.$$

Combining $\frac{\partial \mathcal{L}^{\text{safe}}}{\partial h_{t+1}^s} = 0$ with equation (B.2) yields $\lambda_{ht}^s = 0$. Then, plugging $\frac{\partial \mathcal{L}^{\text{safe}}}{\partial l_t^{f,s}} = 0$ into $\frac{\partial \mathcal{L}^{\text{safe}}}{\partial k_{t+1}^s}$ for λ_{lt}^s, we get

$$E_t\left[R_{t+1}^s\right]Q_t = E_t\left[\alpha\frac{y_{t+1}^s}{k_{t+1}^s} + (1-\delta)Q_{t+1}\right].$$

Consider the zero-profit condition of the safe firm under all states of nature. Since output function has constant returns to scale,

$$y_{t+1}^s = \frac{\partial y_{t+1}^s}{\partial k_{t+1}^s}k_{t+1}^s + \frac{\partial y_{t+1}^s}{\partial h_{t+1}^s}h_{t+1}^s = \alpha A_{t+1}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^{\alpha-1}k_{t+1}^s + W_{t+1}h_{t+1}^s,$$

where we use equation (B.2) to substitute for W_{t+1} in the last equality. Plugging the expression of y_{t+1}^s into $\pi_{t+1}^s = 0$ and using $Q_t k_{t+1}^s = l_t^{f,s}$, we find that:

$$\alpha A_{t+1}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^{\alpha-1}k_{t+1}^s + (1-\delta)Q_{t+1}k_{t+1}^s - R_{t+1}^s Q_t k_{t+1}^s = 0.$$

Since $k_{t+1}^s > 0$, we can divide by k_{t+1}^s to get

$$R_{t+1}^s Q_t = \alpha A_{t+1}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^{\alpha-1} + (1-\delta)Q_{t+1} \tag{B.3}$$

under all states of nature. This condition implies the first-order condition

$$E_t\left[R_{t+1}^s\right]Q_t = E_t\left[\alpha A_{t+1}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^{\alpha-1} + (1-\delta)Q_{t+1}\right].$$

B.2 Risky Firms

Let π_{t+1}^r denote the revenue of a risky firm in period $t+1$ net of expenses:

$$\pi_{t+1}^r = y_{t+1}^r + (1-\delta)Q_t k_{t+1}^r - W_{t+1}h_{t+1}^r - R_{t+1}^r l_t^{f,r}.$$

In this notation, the problem of the risky firm is to

$$\max_{l_t^{f,r}, k_{t+1}^r} E_t \left\{ \max_{h_{t+1}^r} \pi_{t+1}^r \right\}.$$

The first-order condition for $\max_{h_{t+1}^r} \pi_{t+1}^r$ is $\frac{\partial \pi_{t+1}^r}{\partial h_{t+1}^r} = 0$. It implies that

$$W_{t+1} = \frac{\partial y_{t+1}^r}{\partial h_{t+1}^r} = (1-\alpha) A_{t+1} \left(\frac{k_{t+1}^r}{h_{t+1}^r} \right)^\alpha, \tag{B.4}$$

$$h_{t+1}^r = (1-\alpha) \frac{A_{t+1} \left(k_{t+1}^r\right)^\alpha \left(h_{t+1}^r\right)^{1-\alpha}}{W_{t+1}}. \tag{B.5}$$

Accordingly, the risky firm's Lagrangian is:

$$\begin{aligned}\mathcal{L}^{\text{risky}} =& E_t \left[A_{t+1} \left(k_{t+1}^r\right)^\alpha \left(h_{t+1}^r\right)^{1-\alpha} + \varepsilon_{t+1} k_{t+1}^r + (1-\delta) Q_{t+1} k_{t+1}^r - W_{t+1} h_{t+1}^r - R_{t+1}^r l_t^{f,r} \right] + \\ & \lambda_{ht}^r E_t \left[(1-\alpha) \frac{A_{t+1} \left(k_{t+1}^r\right)^\alpha \left(h_{t+1}^r\right)^{1-\alpha}}{W_{t+1}} - h_{t+1}^r \right] + \lambda_{lt}^r \left(l_t^{f,r} - Q_t k_{t+1}^r \right).\end{aligned}$$

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

$$\begin{aligned}\frac{\partial \mathcal{L}^{\text{risky}}}{\partial l_t^{f,r}} &= -E_t \left[R_{t+1}^r \right] + \lambda_{lt}^r = 0, \\ \frac{\partial \mathcal{L}^{\text{risky}}}{\partial k_{t+1}^r} &= E_t \left[\alpha A_{t+1} \left(\frac{k_{t+1}^r}{h_{t+1}^r} \right)^{\alpha-1} + \varepsilon_{t+1} + (1-\delta) Q_{t+1} \right] + \\ & \quad \lambda_{ht}^r E_t \left[\alpha (1-\alpha) \frac{A_{t+1}}{W_{t+1}} \left(\frac{k_{t+1}^r}{h_{t+1}^r} \right)^{\alpha-1} \right] - \lambda_{lt}^r Q_t = 0, \\ \frac{\partial \mathcal{L}^{\text{risky}}}{\partial h_{t+1}^r} &= (1-\alpha) A_{t+1} \left(\frac{k_{t+1}^r}{h_{t+1}^r} \right)^\alpha - W_{t+1} + \lambda_{ht}^r \left[(1-\alpha)^2 \frac{A_{t+1}}{W_{t+1}} \left(\frac{k_{t+1}^r}{h_{t+1}^r} \right)^\alpha - 1 \right] = 0.\end{aligned}$$

Equation (B.4) together with $\frac{\partial \mathcal{L}^{\text{risky}}}{\partial h_{t+1}^r} = 0$ yield $\lambda_{ht}^r = 0$. Plugging $\frac{\partial \mathcal{L}^{\text{risky}}}{\partial l_t^{f,r}} = 0$ into $\frac{\partial \mathcal{L}^{\text{risky}}}{\partial k_{t+1}^r}$ for λ_{lt}^r, we get

$$E_t \left[R_{t+1}^r \right] Q_t = E_t \left[\alpha A_{t+1} \left(\frac{k_{t+1}^r}{h_{t+1}^r} \right)^{\alpha-1} + (1-\delta) Q_{t+1} + \varepsilon_{t+1} \right].$$

Combining equation (B.1) with equation (B.4):

$$\frac{k^s_{t+1}}{h^s_{t+1}} = \frac{k^r_{t+1}}{h^r_{t+1}} \tag{B.6}$$

under all states of nature. But remember that the first-order condition of the safe firm implies

$$E_t\left[R^s_{t+1}\right]Q_t = E_t\left[\alpha A_{t+1}\left(\frac{k^s_{t+1}}{h^s_{t+1}}\right)^{\alpha-1} + (1-\delta)Q_{t+1}\right].$$

Therefore

$$E_t\left[R^s_{t+1}\right]Q_t = E_t\left[R^s_{t+1}Q_t + \varepsilon_{t+1}\right].$$

Consider the zero-profit condition of the risky firm under all states of nature.

$$\begin{aligned}\pi^r_{t+1} = y^r_{t+1} + (1-\delta)Q_t k^r_{t+1} - W_{t+1}h^r_{t+1} - R^r_{t+1}l^{f,r}_t = \\ y^r_{t+1} + (1-\delta)Q_t k^r_{t+1} - (1-\alpha)A_{t+1}\left(k^r_{t+1}\right)^{\alpha}\left(h^r_{t+1}\right)^{1-\alpha} - R^r_{t+1}l^{f,r}_t = \\ \alpha A_{t+1}\left(k^r_{t+1}\right)^{\alpha}\left(h^r_{t+1}\right)^{1-\alpha} + \varepsilon_{t+1}k^r_{t+1} + (1-\delta)Q_t k^r_{t+1} - R^r_{t+1}l^{f,r}_t = \\ \alpha A_{t+1}\left(\frac{k^r_{t+1}}{h^r_{t+1}}\right)^{\alpha-1}k^r_{t+1} + \varepsilon_{t+1}k^r_{t+1} + (1-\delta)Q_t k^r_{t+1} - R^r_{t+1}l^{f,r}_t = 0,\end{aligned}$$

where we use equation (B.5) to substitute for $W_{t+1}h^r_{t+1}$. Using equation (B.3) together with equation (B.6), we can express

$$\alpha A_{t+1}\left(\frac{k^r_{t+1}}{h^r_{t+1}}\right)^{\alpha-1} = R^s_{t+1}Q_t - (1-\delta)Q_{t+1},$$

that holds under all states of nature. Plugging it into the zero-profit condition and using $Q_t k^r_{t+1} = l^{f,r}_t$, we find that:

$$R^s_{t+1}Q_t k^r_{t+1} - (1-\delta)Q_{t+1}k^r_{t+1} + \varepsilon_{t+1}k^r_{t+1} + (1-\delta)Q_t k^r_{t+1} - R^r_{t+1}Q_t k^r_{t+1} = 0.$$

Since $k^r_{t+1} > 0$, we can divide by k^r_{t+1} to get

$$R^r_{t+1}Q_t = R^s_{t+1}Q_t + \varepsilon_{t+1}$$

under all states of nature. This condition implies

$$E_t\left[R^r_{t+1}\right]Q_t = E_t\left[R^s_{t+1}Q_t + \varepsilon_{t+1}\right].$$

B.3 Aggregating across firms

Here we show that we can aggregate individual firms into two representative firms. Let denote $k^i_{j,t}$ the capital chosen by firm i that is financed by borrowing from bank j. Both i and j lie within the continuum of measure 1 of banks and firms, respectively. In this notation, the equation (B.6) is written as

$$\frac{k^i_{j,t+1}}{h^i_{j,t+1}} = \frac{k_{t+1}}{h_{t+1}}, \tag{B.7}$$

for all $j \in [0,1]$ and $i \in [0,1]$. Each firm chooses the same capital-to-labor ratio independently of the type of bank it borrows from.

Notice is that σ_t is the fraction of risky firms at date t; the remaining fraction $1-\sigma_t$ of firms are safe firms. Let's index firms as follows: firm j_1, with $j_1 \in [0, \sigma_t]$, can only access a risky technology subject to both aggregate and idiosyncratic shocks; firm j_2, with $j_2 \in [\sigma_t, 1]$ has access to a safe production technology subject to aggregate shocks only. Since there are no equilibria with $\underline{\sigma} < \sigma_t < \bar{\sigma}$, the fraction of risky firms is linked to the fraction of banks with risky portfolios as follows:

$$\sigma_t = (1-\mu_t)\,\underline{\sigma} + \mu_t \bar{\sigma}.$$

Define the following objects: Let $K^s_{s,t+1} = \int_{\sigma_t}^1 \int_{\mu_t}^1 k^i_{j,t+1} dj di$ 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 $K^s_{r,t+1} = \int_{\sigma_t}^1 \int_0^{\mu_t} k^i_{j,t+1} dj di$ 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. We let K^s_{t+1} denote the total capital allocated to the safe technology. Thus,

$$K^s_{t+1} = \int_{\sigma_t}^1 \int_0^1 k^i_{j,t+1} dj di = K^s_{s,t+1} + K^s_{r,t+1},$$

Let $K^r_{s,t+1} = \int_0^{\sigma_t} \int_{\mu_t}^1 k^i_{j,t+1} dj di$ 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 $K^r_{r,t+1} = \int_0^{\sigma_t} \int_0^{\mu_t} k^i_{j,t+1} dj di$ 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. We let K^r_{t+1} denote the total capital allocated to the risky technology. Thus,

$$K^r_{t+1} = \int_0^{\sigma_t} \int_0^1 k^i_{j,t+1} dj di = K^r_{s,t+1} + K^r_{r,t+1},$$

The same upper and lower case notation applies to labor, i.e. $H^s_{s,t+1} = \int_{\sigma_t}^1 \int_{\mu_t}^1 h^i_{j,t+1} dj di$; $H^s_{r,t+1} = \int_{\sigma_t}^1 \int_0^{\mu_t} h^i_{j,t+1} dj di$; $H^r_{s,t+1} = \int_0^{\sigma_t} \int_{\mu_t}^1 h^i_{j,t+1} dj di$; $H^r_{r,t+1} = \int_0^{\sigma_t} \int_0^{\mu_t} h^i_{j,t+1} dj di$.

Safe representative firm produces:

$$Y^s_t = \int\limits_{\sigma_{t-1}}^{1} \int\limits_{0}^{1} A_t \left(k^i_{j,t}\right)^{\alpha} \left(h^i_{j,t}\right)^{1-\alpha} dj di = \int\limits_{\sigma_{t-1}}^{1} \int\limits_{0}^{1} F\left(k^i_{j,t}, h^i_{j,t}\right) dj di =$$

Using that the technology has Constant Returns to Scale:

$$= \int\limits_{\sigma_{t-1}}^{1} \int\limits_{0}^{1} \left[F_{k^i_{j,t}}\left(k^i_{j,t}, h^i_{j,t}\right) k^i_{j,t} + F_{h^i_{j,t}}\left(k^i_{j,t}, h^i_{j,t}\right) h^i_{j,t}\right] dj di =$$

where $F_{k^i_{j,t}}\left(k^i_{j,t}, h^i_{j,t}\right)$ and $F_{h^i_{j,t}}\left(k^i_{j,t}, h^i_{j,t}\right)$ denote the partial derivative of $F\left(k^i_{j,t}, h^i_{j,t}\right)$ with respect to $k^i_{j,t}$ and $h^i_{j,t}$, respectively. Since these partial derivatives are homogeneous of degree zero, we can express them in term of capital-labor ratio, i.e.

$$= \int\limits_{\sigma_{t-1}}^{1} \int\limits_{0}^{1} \left[f_{k^i_{j,t}}\left(\frac{k^i_{j,t}}{h^i_{j,t}}\right) k^i_{j,t} + f_{h^i_{j,t}}\left(\frac{k^i_{j,t}}{h^i_{j,t}}\right) h^i_{j,t}\right] dj di = \text{Plugging equation (B.7)} =$$

$$= \int\limits_{\sigma_{t-1}}^{1} \int\limits_{0}^{1} \left[f_{k_t}\left(\frac{k_t}{h_t}\right) k^i_{j,t} + f_{h_t}\left(\frac{k_t}{h_t}\right) h^i_{j,t}\right] dj di =$$

$$f_{k_t}\left(\frac{k_t}{h_t}\right)\left[\int\limits_{\sigma_t}^{1} \int\limits_{0}^{1} k^i_{j,t} dj di\right] + f_{h_t}\left(\frac{k_t}{h_t}\right)\left[\int\limits_{\sigma_t}^{1} \int\limits_{0}^{1} h^i_{j,t} dj di\right] = f_{k_t}\left(\frac{k_t}{h_t}\right) K^s_t + f_{h_t}\left(\frac{k_t}{h_t}\right) H^s_t =$$

Since $\frac{K^s_{s,t}}{H^s_{s,t}} = \frac{K^s_{r,t}}{H^s_{r,t}} = \frac{k_t}{h_t}$,then $\frac{K^s_t}{H^s_t}\frac{h_t}{k_t} = \left(\frac{K^s_{s,t}+K^s_{r,t}}{H^s_{s,t}+H^s_{r,t}}\right)\frac{H^s_{r,t}}{K^s_{r,t}} = 1$. Therefore $\frac{K^s_t}{H^s_t} = \frac{k_t}{h_t}$.

$$= f_{K^s_t}\left(\frac{K^s_t}{H^s_t}\right) K^s_t + f_{H^s_t}\left(\frac{K^s_t}{H^s_t}\right) H^s_t = A_t \left(K^s_t\right)^{\alpha} \left(H^s_t\right)^{1-\alpha}.$$

Risky representative firm:

$$Y^r_t = \int\limits_{0}^{\sigma_{t-1}} \int\limits_{0}^{1} \left[A_t \left(k^i_{j,t}\right)^{\alpha} \left(h^i_{j,t}\right)^{1-\alpha} + \varepsilon^i_{j,t} k^i_{j,t}\right] dj di = \int\limits_{0}^{\sigma_{t-1}} \int\limits_{0}^{1} F\left(k^i_{j,t}, h^i_{j,t}\right) dj di + \int\limits_{0}^{\sigma_{t-1}} \int\limits_{0}^{1} \varepsilon^i_{j,t} k^i_{j,t} dj di$$

Note that the similar steps described above apply to the first term in the summation, so that $\int_0^{\sigma_{t-1}} \int_0^1 F\left(k^i_{j,t}, h^i_{j,t}\right) dj di = A_t \left(K^r_t\right)^{\alpha} \left(H^r_t\right)^{1-\alpha}$. To express the second term, notice that $\int_0^{\sigma_{t-1}} \int_0^1 \varepsilon_{j,t}^i k_{j,t}^i dj di = -\xi$. Moreover since each risky firm solves the same maximization problem, it chooses the same amount of capital independently of the type of bank it borrows from. Therefore, $\int_0^{\sigma_{t-1}} \int_0^1 \varepsilon_{j,t}^i k_{j,t}^i dj di = -\xi K_t^r$. Hence,

$$Y_t^r = A_t \left(K_t^r\right)^\alpha \left(H_t^r\right)^{1-\alpha} - \xi K_t^r.$$

C The Government

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

C.1 Baseline: No linear cost of banking

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

$$- \left[\int_{-\infty}^{\infty} \left(\left(R_t^l + \frac{\sigma_{t-1}\varepsilon_t}{Q_{t-1}}\right) L_{t-1} - R_{t-1}^d D_{t-1}\right) \mathrm{d}G(\varepsilon_t) - \right.$$

$$\left. \int_{\left(\frac{R_{t-1}^d D_{t-1}}{\sigma_{t-1} L_{t-1}} - \frac{R_t^s}{\sigma_{t-1}}\right) Q_{t-1}}^{\infty} \left(\left(R_t^s + \frac{\sigma_{t-1}\varepsilon_t}{Q_{t-1}}\right) L_{t-1} - R_{t-1}^d D_{t-1}\right) \mathrm{d}G(\varepsilon_t) \right] =$$

Note that the first term equals $\left(R_t^s - \frac{\sigma_{t-1}\xi}{Q_{t-1}}\right) L_{t-1} + R_{t-1}^d D_{t-1}$ in the square bracket. We have already calculated the second term. Therefore,

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

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

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

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

C.2 Linear Cost of Banking: Tax

The tax that accounts for the cost of banking is described as follows:

$$T_t = -\int_{-\infty}^{\left(\frac{R^d_{t-1}d_{t-1}}{\sigma_{t-1}l_{t-1}} - \frac{R^s_t - f}{\sigma_{t-1}}\right)Q_{t-1}} \left(\left(R^s_t + \frac{\sigma_{t-1}\varepsilon_t}{Q_{t-1}} - f\right) l_{t-1} - R^d_{t-1}d_{t-1}\right) \mathrm{d}G(\varepsilon_t) =$$

$$\frac{\sigma_{t-1}l_{t-1}}{Q_{t-1}}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(f+R^d_{t-1}(1-\gamma_{t-1})-R^s_t\right)Q_{t-1}+\xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)^2} - \left(R^l_t - \frac{\sigma_{t-1}\xi}{Q_{t-1}} - f\right) l_{t-1} + R^d_{t-1}d_{t-1} +$$

$$\frac{1}{2}l_{t-1}\left(R^s_t - \frac{\sigma_{t-1}\xi}{Q_{t-1}} - (1-\gamma_{t-1})R^d_{t-1} - f\right)\left[1 - \mathrm{erf}\left(\frac{\left(f + R^d_{t-1}(1-\gamma_{t-1}) - R^s_t\right)Q_{t-1} + \xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)\right] =$$

$$\frac{\sigma_{t-1}l_{t-1}}{Q_{t-1}}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(f+R^d_{t-1}(1-\gamma_{t-1})-R^s_t\right)Q_{t-1}+\xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)^2} -$$

$$\frac{1}{2}\left(R^s_t l_{t-1} - \frac{\sigma_{t-1}\xi}{Q_{t-1}}l_{t-1} - R^d_{t-1}d_{t-1} - f l_{t-1}\right)\left[1 + \mathrm{erf}\left(\frac{\left(f + R^d_{t-1}(1-\gamma_{t-1}) - R^s_t\right)Q_{t-1} + \xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)\right].$$

D Choice of Risk

This appendix shows a proof that the expected dividends function of banks is convex in the risk parameter σ_t. This result guarantees that banks choose either the maximum risk, $\bar{\sigma}$, or the minimum risk, $\underline{\sigma}$, to maximize their profits, so all the intermediate values of σ_t, which may result from the first-order conditions with respect to σ_t, are not optimal.

We 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}, with non-positive mean, so our example of a Normal distribution considered in the analysis is not a special case which can drive our results. It is used for expositional reasons and quantitative work.

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 $-\xi$ ($\xi > 0$). ε is independent of the aggregate shock. The aggregate shock does not depend on the choice of σ_t.*

Note that we do not restrict the analysis to the bounded support[18], so $\underline{\varepsilon}$ and $\bar{\varepsilon}$ can take $-\infty$ and $+\infty$, respectively. Note that G_ε need not be continuous.

Let $\hat{\varepsilon}(\sigma_t, R^s_{t+1}) \equiv \left(\frac{R^d_t d_t}{\sigma_t l_t} - \frac{R^l_{t+1}}{\sigma_t}\right) Q_t = \frac{R^d_t(1-\gamma_t) - R^s_{t+1}}{\sigma_t} Q_t$, where the latter equation uses the result that the capital requirement constraint always binds. Therefore, $\left(R^s_{t+1} + \sigma_t \frac{\hat{\varepsilon}(\sigma_t)}{Q_t}\right) l_t -$

[18] Unbounded support is more relevant if we consider aggregate risk

$R_t^d d_t = 0$. Let $\pi(\sigma_t,\ R^s_{t+1}) = E_\varepsilon\left[\left(\left(R^s_{t+1}+\frac{\sigma_t\varepsilon}{Q_t}\right)l_t - R_t^d d_t\right)^+\right]$ be a function of expected dividends (taken over the idiosyncratic shock only) under some realization of R^s_{t+1} which is considered to be fixed in this function. To account for the aggregate uncertainty, R^s_{t+1} needs to be a random variable. Therefore, expected dividends taken into account both idiosyncratic and aggregate uncertainty are

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

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

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

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

Note that in the derivations above we express $\left(R^s_{t+1}+\frac{\sigma_t\varepsilon}{Q_t}\right)l_t - R_t^d d_t$ in terms of $\hat{\varepsilon}(\sigma_t,\ R^s_{t+1})$ and ε using the definition of $\hat{\varepsilon}(\sigma_t,\ R^s_{t+1})$.

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 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. $H(\sigma_t) \equiv l_t E_t\left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t)} \left(\hat{\varepsilon}(\sigma_t,\ R^s_{t+1}) - \varepsilon\right)dG_\varepsilon\right]\sigma_t$ *is convex in* σ_t*:*

Proof. Steps of the proof: □

1. Define $h(\sigma_t,\ R^s_{t+1}) \equiv \sigma_t\left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t,\, R^s_{t+1})} \left(\hat{\varepsilon}(\sigma_t,\ R^s_{t+1}) - \varepsilon\right)dG_\varepsilon\right]$ in which the aggregate uncertainty is taken off. Consider 3 cases:

   (a) Realization of R^s_{t+1} is such that $\hat{\varepsilon}(\sigma_t,\ R^s_{t+1}) = \frac{R_t^d(1-\gamma_t)-R^s_{t+1}}{\sigma_t} > 0$, so $R^s_{t+1} < R_t^d\left(1-\gamma_t\right)$,

(b) Realization of R^s_{t+1} is such that $\hat{\varepsilon}(\sigma_t,\, R^l_{t+1}) = \frac{R^d_t(1-\gamma_t)-R^s_{t+1}}{\sigma_t} < 0$, so $R^s_{t+1} > R^d_t\,(1-\gamma_t)\,,$

(c) Realization of R^s_{t+1} is such that $\hat{\varepsilon}(\sigma_t,\, R^l_{t+1}) = \frac{R^d_t(1-\gamma_t)-R^s_{t+1}}{\sigma_t} = 0$, so $R^s_{t+1} = R^d_t\,(1-\gamma_t)\,,$

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

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

Let's 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 $\hat{\varepsilon}_i = \hat{\varepsilon}(\sigma_{it},\, R^s_{t+1}) \equiv \frac{R^d_t(1-\gamma_t)-R^s_{t+1}}{\sigma_{it}} Q_t$, for $i = 1,\, 2,\, \lambda$.

   (a) $R^s_{t+1} < R^d_t\,(1-\gamma_t)$: it implies that $\hat{\varepsilon}_2 < \hat{\varepsilon}_\lambda < \hat{\varepsilon}_1$,

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

   where the inequality sign comes from $\int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\varepsilon - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \leq \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \hat{\varepsilon}_\lambda)\, dG_\varepsilon$ and $\int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \leq \int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, dG_\varepsilon$. Substituting for the definitions of $h(\sigma_{1t}) =$

$\sigma_{1t} \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon$ and $h(\sigma_{2t}) = \sigma_{2t} \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon$, we 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 we use that $\sigma_{1t} = l_t \left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) = \sigma_{2t}\hat{\varepsilon}_2 = \sigma_{\lambda t}\hat{\varepsilon}_\lambda$ in the last equality. So,

$$\lambda \sigma_{1t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1) + (1-\lambda)\sigma_{2t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2) = \hat{\varepsilon}_\lambda (\lambda \sigma_{1t} + (1-\lambda)\sigma_{2t}) - \left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) (\lambda + (1-\lambda)) = \sigma_{\lambda t}\hat{\varepsilon}_\lambda - \left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) = \left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) - \left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) = 0.$$

Therefore, $h(\sigma_t)$ is convex in σ_t for $R_{t+1}^s < R_t^d (1-\gamma_t)$.

(b) $R_{t+1}^s > R_t^d (1-\gamma_t)$: it implies that $\hat{\varepsilon}_1 < \hat{\varepsilon}_\lambda < \hat{\varepsilon}_2$

$$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\},$$

where the inequality sign comes from $\int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \leq \int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, dG_\varepsilon$ and $\int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\varepsilon - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \leq \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \hat{\varepsilon}_\lambda)\, dG_\varepsilon$. Substituting for the definitions of $h(\sigma_{1t}) =$

$\sigma_{1t} \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon$ and $h(\sigma_{2t}) = \sigma_{2t} \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon$, we 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 the same reasoning employed in the previous case. Therefore, $h(\sigma_t)$ is convex in σ_t for $R^s_{t+1} > R^d_t(1-\gamma_t)$.

(c) $R^s_{t+1} = R^d_t(1-\gamma_t)$. Hence, $\hat{\varepsilon}(\sigma_t) = 0$ and

$$h(\sigma_t) = \sigma_t \left[\int_{\underline{\varepsilon}}^{0} (0 - \varepsilon)\, dG_\varepsilon\right],$$

which is linear in σ_t

2. We found in 1 that $h(\sigma_t,\, R^s_{t+1})$ is convex in σ_t for each $R^s_{t+1} \in \mathbb{R}$. Consider $P(\omega) \geq 0$ for each $R^l_{t+1}(\omega) \in \mathbb{R}$. Then the following function[19]:

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

is convex in σ_t. It follows directly from the linearity of the expectation operator which puts a non-negative weight on every realization of R^s_{t+1} and the fact that the sum of convex functions is a convex function. Therefore, $\Pi(\sigma_t)$ is convex in σ_t. □

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

E Equilibrium Conditions

For $\forall i \in [s, r]$:

$$(C_t - \kappa C_{t-1})^{-\varsigma_c} - \beta\kappa E_t (C_{t+1} - \kappa C_t)^{-\varsigma_c} - \lambda_{ct} = 0 \tag{E.1}$$

$$\varsigma_0 D_t^{-\varsigma_d} - \lambda_{ct} + E_t \beta \lambda_{ct+1} R_t^d = 0, \tag{E.2}$$

$$-\lambda_{ct} + E_t \beta \lambda_{ct+1} R_{t+1}^{e,s} + \zeta_t^s = 0, \tag{E.3}$$

$$-\lambda_{ct} + E_t \beta \lambda_{ct+1} R_{t+1}^{e,r} + \zeta_t^r = 0, \tag{E.4}$$

$$\zeta_t^s E_t^s = 0, \tag{E.5}$$

$$\zeta_t^r E_t^r = 0 \tag{E.6}$$

$$\begin{aligned} \gamma_t - \chi_{2t}^i = E_t \Bigg\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \Bigg[\frac{\sigma_t^i}{Q_t} \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) Q_t + \xi \sigma_t^i}{\sigma_t^i \sqrt{2}\tau} \right)^2} + \\ \frac{1}{2} \left(R_{t+1}^s - \frac{\sigma_t^i \xi}{Q_t} - R_t^d \right) \left[1 - \mathrm{erf} \left(\frac{\left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) Q_t + \xi \sigma_t^i}{\sigma_t^i \sqrt{2}\tau} \right) \right] \Bigg] \Bigg\}, \end{aligned} \tag{E.7}$$

$$\begin{aligned} R_{t+1}^{e,i} = \frac{1}{\gamma_t} \Bigg\{ \frac{\sigma_t^i}{Q_t} \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) Q_t + \xi \sigma_t^i}{\sigma_t^i \sqrt{2}\tau} \right)^2} + \\ \frac{1}{2} \left(R_{t+1}^s - \frac{\sigma_t^i \xi}{Q_t} - R_t^d \right) \left[1 - \mathrm{erf} \left(\frac{\left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) Q_t + \xi \sigma_t^i}{\sigma_t^i \sqrt{2}\tau} \right) \right] \Bigg\}, \end{aligned} \tag{E.8}$$

$$\chi_{2t}^i l_t^i = 0, \tag{E.9}$$

$$\sigma^s = \underline{\sigma}, \tag{E.10}$$

$$\sigma^r = \bar{\sigma}, \tag{E.11}$$

$$l_t^i = d_t^i + e_t^i, \tag{E.12}$$

$$e_t^i = \gamma_t l_t^i, \tag{E.13}$$

$$\Omega(\sigma_t^i;\, l_t^i,\, d_t^i,\, e_t^i) = E_t \left[\beta \tfrac{\lambda_{ct+1}}{\lambda_{ct}} R_{t+1}^{e,i} e_t^i \right], \tag{E.14}$$

$$\mu_t = \tfrac{E_t^r}{E_t^s + E_t^r}, \tag{E.15}$$

$$L_t^s = (1-\mu_t)\, l_t^s, \tag{E.16}$$

$$L_t^r = \mu_t l_t^r, \tag{E.17}$$

$$E_t^i = \gamma_t L_t^i, \tag{E.18}$$

$$L_t^i = D_t^i + E_t^i, \tag{E.19}$$

$$D_t = D_t^s + D_t^r, \tag{E.20}$$

$$Y_t^s = A_t \left(K_t^s\right)^{\alpha} \left(H_t^s\right)^{1-\alpha}, \tag{E.21}$$

$$Y_t^r = A_t \left(K_t^r\right)^{\alpha} \left(H_t^r\right)^{1-\alpha} - \xi K_t^r, \tag{E.22}$$

$$Q_t K_{t+1}^s = \left(1-\underline{\sigma}\right) L_t^s + \left(1-\bar{\sigma}\right) L_t^r, \tag{E.23}$$

$$Q_t K_{t+1}^r = \underline{\sigma} L_t^s + \bar{\sigma} L_t^r, \tag{E.24}$$

$$W_t = (1-\alpha)\, \tfrac{Y_t^s}{H_t^s}, \tag{E.25}$$

$$R_t^s = \tfrac{\alpha A_t}{Q_t} \left(\tfrac{K_t^s}{H_t^s}\right)^{\alpha-1} + (1-\delta)\, \tfrac{Q_{t+1}}{Q_t}, \tag{E.26}$$

$$R_t^r = R_t^s + \tfrac{\varepsilon_t}{Q_{t-1}}, \tag{E.27}$$

$$\tfrac{K_t^s}{H_t^s} = \tfrac{K_t^r}{H_t^r}, \tag{E.28}$$

$$H_t^s + H_t^r = 1, \tag{E.29}$$

$$K_t = K_t^s + K_t^r, \tag{E.30}$$

$$K_{t+1} = I_t + (1-\delta) K_t, \tag{E.31}$$

$$I_t = \eta_t \left[1 - \tfrac{\phi}{2}\left(\tfrac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] I_t^g, \tag{E.32}$$

$$\begin{aligned} \eta_t Q_t \left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] - \eta_t Q_t \phi \left(\frac{I_t^g}{I_{t-1}^g} - 1\right) \frac{I_t^g}{I_{t-1}^g} - 1 + \\ \eta_{t+1} \psi_{t,t+1} Q_{t+1} \phi \left(\frac{I_{t+1}^g}{I_t^g} - 1\right) \frac{I_{t+1}^g}{\left(I_t^g\right)^2} I_{t+1}^g = 0, \end{aligned} \tag{E.33}$$

$$Y_t^s + Y_t^r = C_t + I_t^g, \tag{E.34}$$

$$\begin{aligned} T_t = L_{t-1} \Bigg\{ & \frac{\sigma_{t-1}}{Q_{t-1}} \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_{t-1}^d (1-\gamma_{t-1}) - R_t^s\right) Q_{t-1} + \xi \sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)^2} - \\ & \frac{1}{2}\left(R_t^s - R_{t-1}^d \left(1-\gamma_{t-1}\right) - \frac{\xi \sigma_{t-1}}{Q_{t-1}}\right) \left[1 + \operatorname{erf}\left(\frac{\left(R_{t-1}^d \left(1-\gamma_{t-1}\right) - R_t^s\right) Q_{t-1} + \xi \sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)\right] \Bigg\}. \end{aligned} \tag{E.35}$$

F Discussion of the Excessive Risk-Taking Mechanism

Following our the result derived earlier, we can express the erf function in terms of the share of non-defaulted deposits of the representative bank and then decompose the expected dividend into two components:

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

where

$$\left[\omega_1 + \omega_2\right] = \left[\underbrace{\left(R^s_{t+1} - R^d_t\left(1-\gamma_t\right) - \frac{\xi\sigma_t}{Q_t}\right)\underbrace{\left(1-G(\varepsilon^*_{t+1})\right)}_{\text{non-defaulted}}}_{\substack{\omega_1 \equiv \text{ returns from a loan} \\ \text{portfolio with riskiness } \sigma_t}} + \underbrace{\left(\frac{\sigma_t}{Q_t}\right)\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\varepsilon^*_{t+1}+\xi}{\tau\sqrt{2}}\right)^2}}_{\substack{\omega_2 \equiv \text{ bonus from} \\ \text{projects volatility}}}\right],$$

and the cutoff point ε^*_{t+1} is defined by $R^d_t\left(1-\gamma_t\right)Q_t - R^s_{t+1}Q_t = \sigma_t\varepsilon^*_{t+1}$.

The first component, ω_1, distinguishes loan returns of riskiness σ_t controlling for the variance of idiosyncratic shock (when τ is taken as given). The bank trades off the benefits from limited liability and deposit insurance with a smaller profitability of riskier projects. The term $\frac{\xi\sigma_t}{Q_t}$ reflects, in expectation, the reduction of loan returns for the bank holding σ_t share of risky projects. The bank receives net income on loans, $R^s_{t+1} - R^d_t\left(1-\gamma_t\right) - \frac{\xi\sigma_t}{Q_t}$, if it does not default on deposits which happens with probability $1 - G(\varepsilon^*_{t+1})$. If the bank defaults, it gets zero, i.e. $0 \cdot G(\varepsilon^*_{t+1})$ which is not shown in the expression explicitly.

The second counterpart of the above decomposition, ω_2, comprises the extra effect of σ_t on expected dividends that comes from more dispersed returns from projects. In fact, ω_2 is strictly increasing in τ: the bank views projects as a call option the value of which rises with volatility associated with higher upside. Limited liability bounds the payoff to zero in the worst case scenario.

Risk-taking incentives depend on the difference between returns on safe loans and returns on deposits. Table 11 illustrates the effects of greater risk-taking on two components of dividends for each realization of the aggregate returns. We map aggregate returns into states of nature and consider two cases depending on the sign of ε^*_{t+1}. The aggregate returns influence the value of the shield of limited liability. Risk amplifies the effect of the idiosyncratic shock. So, in every state of nature, the bank's choice of risk is determined by the expected effect of the idiosyncratic shock on the value of the shield of limited liability and returns on loans. The up-turn arrow, $\Uparrow$, indicates that greater risk-taking increases the corresponding component of bank's dividends. The down-turn arrow, $\Downarrow$, means that the corresponding component of bank's dividends decreases with greater risk-taking. Two arrows turned in the opposite directions, $\Uparrow\Downarrow$, signify that the effect of greater risk-taking is undetermined and depends the parameterization.

Table 11: Illustrating the Effects of Higher Risk on Dividends.

<table>
<tr><th rowspan="2">States of nature where</th><th colspan="2">Effects on ω_1</th><th rowspan="2">Effects on ω_2</th></tr>
<tr><th>$R^l_{t+1} - R^d_t(1-\gamma_t) - \frac{\xi\sigma_t}{Q_t}$</th><th>$1 - G(\varepsilon^*_{t+1})$</th></tr>
<tr><td>$R^l_{t+1} < R^d_t(1-\gamma_t) \quad \Leftrightarrow \quad \varepsilon^*_{t+1} > 0$</td><td>$\Downarrow$</td><td>$\Uparrow$</td><td>$\Uparrow$</td></tr>
<tr><td rowspan="2">$R^l_{t+1} > R^d_t(1-\gamma_t) \quad \Leftrightarrow \quad \varepsilon^*_{t+1} < 0$</td><td rowspan="2">$\Downarrow$</td><td rowspan="2">$\Downarrow$</td><td>if $\varepsilon^*_{t+1} > -\xi$, then $\Uparrow\Downarrow$</td></tr>
<tr><td>if $\varepsilon^*_{t+1} \leqslant -\xi$, then $\Uparrow$</td></tr>
</table>

First, $\varepsilon^*_{t+1} > 0$ indicates that the bank makes losses on safe loans. It happens in those states of nature where the net income from the zero-risk portfolio is negative, so the bank is behind the shield of limited liability. By accepting more risk, the bank is more likely to get a positive net return under a favorable realization of the idiosyncratic shock as risk acts like a leverage on the size of the shock. Therefore, $1 - G(\varepsilon^*_{t+1})$ rises. This balances with smaller returns on a portfolio with more risky loans, i.e. $R^s_{t+1} - R^d_t(1-\gamma_t) - \frac{\xi\sigma_t}{Q_t}$ goes down. Similarly, gambling on more dispersed returns allows the bank to move away from a zero return that comes from the limited liability to some positive return that is accompanied by less frequent defaults. So, the effect of σ_t on expected dividends from ω_2 is positive.

Second, $\varepsilon^*_{t+1} < 0$ shows that the bank makes positive profits on safe loans. The bank is more likely to default when it takes on more risk because any negative idiosyncratic shock would be amplified by risk. The bank internalizes that riskier projects are less profitable. Therefore, the overall effect of greater risk on ω_1 is negative when $\varepsilon^*_{t+1} < 0$.

Then consider the bonus from projects volatility. If $-\xi < \varepsilon^*_{t+1} < 0$, there are two contrasting forces. On the one hand, the bank always benefits from limited liability that makes the variance of projects returns attractive. On the other hand, the bank is more concerned about (and more vulnerable to) the variability of returns in the situation when taking on more risk would result in zero payoff instead of some positive payoff achieved by smaller risk. It occurs when $-\xi < \varepsilon^*_{t+1} < 0$. In these states of nature, the bank requires greater than average realization of the idiosyncratic shock in order to get a positive return. Call this type of shock a good idiosyncratic shock. This shock happens with probability smaller than 0.5. Define a bad idiosyncratic shock as a complement to a good idiosyncratic shock. An increase in risk increases the profits under a good shock. It captures the benefits from greater upside. At the same time, an increase in risk makes it more likely to get a bad shock. The bank trades off marginal profits coming from a good shock with marginal losses coming from the reduction of profits due to more defaults. Since the probability of the latter is greater than the probability of the former, the losses from defaults can dominate the benefits from greater volatility. This force goes in the opposite direction when $\varepsilon^*_{t+1} \leqslant -\xi$. The difference is that here the bank is more likely to get a good shock than a bad shock. Therefore, the bank puts more weight on the benefits from risk-taking than on its costs. It is verified mathematically that the effects of σ_t on ω_2 is unambiguously positive when $\varepsilon^*_{t+1} \leqslant -\xi$.

In sum, we find that net returns on safe loans, $R^s_{t+1} - R^d_t (1 - \gamma_t)$, is the main driver for the bank's choice of risk. In the partial-equilibrium setting, we differentiate between three cases that characterize incentives for risk-taking.

First, $R^s_{t+1} < R^d_t (1 - \gamma_t)$ applies to the states of nature where a relatively large negative aggregate shock is realized. Two forces against the one that seems to be of lesser relevance make the bank benefit most from taking risk. Second, $-\xi < R^d_t (1 - \gamma_t) - R^s_{t+1} < 0$ applies to the states of nature where intermediate values (not too large and not too small) of either negative or positive aggregate shock are realized. There are more forces that lower incentives for risk. Third, $R^d_t (1 - \gamma_t) - R^s_{t+1} < -\xi$ applies to the states of nature where a positive aggregate shock of a larger size is realized. Interestingly, there is a force associated with the bonus from projects volatility that makes it possible for the bank to increase risk. The choice of risk depends on the strength of that force, ω_2, relative to the negative exposure of returns from a loan portfolio to risk, ω_1. It still remains a quantitative question to find out how risk-taking is determined in the general equilibrium set-up.

Capital requirements affect risk-taking through a change in ε^*_{t+1}. When γ_t increases, ε^*_{t+1} falls. It means that the bank will be more likely to find itself in the states of nature where ε^*_{t+1} is negative. It forces the bank to keep more skin in the game, make the shield of limited liability less attractive and prevent the switch into financing risky projects.

G Calibration of τ

To calibrate the variance of the idiosyncratic shock τ, we 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

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

respectively. The parameter σ_{RF} is needed to distill the exposure of banks (versus other financial intermediaries) to the risk arising in the leveraged loan market. It captures the fact that a certain fraction of leveraged loans is held by the nonbank sector which we do not model here. The risky bank that finances the maximum share of risky projects earns

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

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

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

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

$$Var_t\left(R^r_{t+1}\right) = Var_t\left(R^s_{t+1}\left(1 + \frac{B_t}{K_{t+1}}\right) - \frac{B_t}{K_{t+1}} R^B_t\right)$$

to find the variance of the idiosyncratic shock that matches $\frac{\text{Debt}}{\text{EBITDA}} = 6$. Note that

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

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

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

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

We can calculate the conditional variance of Q_{t+1} by picking up its process from the optimization problem of capital producers. However, our approach is meant to be suggestive, and we equate the conditional variances of Q_{t+1} and the aggregate shock. The covariance term is expected to be positive, but we drop it in our calculation because the terms that multiply the covariance 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 banks (where the remaining fraction is held by nonbanks). We choose $\sigma_{\mathrm{RF}} = 45\%$ from the Shared National Credit Report issued by the Fed, OCC, and FDIC.

H Robustness Checks

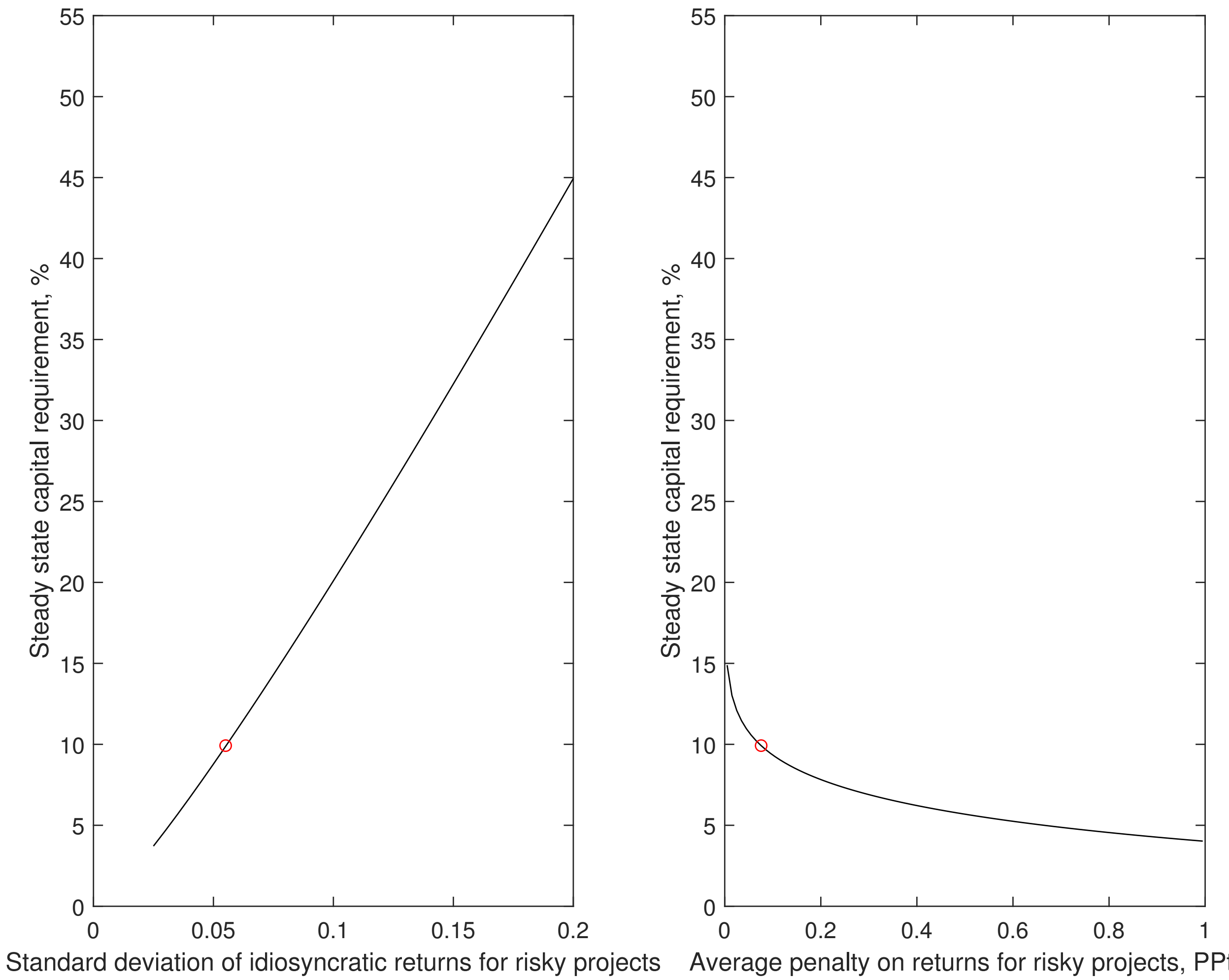


Figure 7: Robustness Checks.