A Static Capital Buffer is Hard To Beat*

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

First Draft: September 1, 2021;[‖] this Draft February 15, 2024

Abstract

In a model with endogenous risk taking, deposit insurance and limited liability may lead banks to make risky loans that are socially inefficient. Capital requirements can prevent excessive risk taking at the cost of reducing liquidity-producing bank deposits. The all-knowing Ramsey planner will set capital requirements pro-, counter-, or a-cyclically depending on the origin of each shock. However, the Ramsey response to a constellation of shocks is not implementable. Simple rules that respond to cyclical conditions—in line with Basel III guidance—perform poorly, whereas a small, static capital buffer does nearly as well as the optimal Ramsey policy.

JEL Classification: C54, E13, G21

Keywords: banks, capital requirements, endogenous risk taking, crises.

*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, Aaron Pancost, and Skander Van den Heuvel 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 real returns on safe assets followed in the wake of the global financial crisis of 2008, and this trend is expected to continue after the Federal Reserve normalizes policy following the Covid inflation.[1] These low returns have raised concerns that financial intermediaries will once again be tempted to reach for higher yields by taking excessive (or socially inefficient) risks. The risk-taking behavior that we have in mind is epitomized by the emergence of NINJA loans in the subprime mortgage market leading up to the Great Recession, and by the booming of leveraged loans in its aftermath.

To study these concerns, we develop a dynamic macroeconomic model in which limited liability and deposit insurance provide incentives for a bank to shift from safe assets to risky assets in its portfolio of loans.[2] More specifically, and following Van den Heuvel (2008), our banks can lend to safe firms or risky firms. As in that work, risky firms are exposed to an idiosyncratic shock with negative expected value. A profit maximizing bank could fund a firm with negative expected value only because limited liability shields it from downside risk and because deposit insurance takes away the incentives of depositors to monitor the activities of banks. An important extension in our model is that both safe and risky firms face aggregate shocks, allowing the model to capture aggregate fluctuations. In response to shocks, if capital requirements are not sufficiently high, financing risky loans can temporarily become attractive and lead to banking crises in which some banks fail and deposit insurance bails out depositors.[3]

Our contribution is to calculate the optimal Ramsey policy for capital requirements and compare it to simple and implementable rules in the more realistic situation where a constellation of shocks bombards the economy at the same time. We show that the optimal

[1]Estimates of the natural rate (or what the Fed calls r^*) are currently in the range of 50 to 100 basis points; see John C. Williams (2023) and the references therein.

[2]We do not analyze the optimality of either limited liability or bank deposit insurance; we simply take them as given constraints on the Ramsey planner. Our model would not be adequate for a discussion of deposit insurance since we exclude the possibility of bank runs.

[3]This is another important extension relative to Van den Heuvel (2008), who only considers excessive risk taking as an off-equilibrium outcome, ruled out by sufficiently high capital requirements. While that work came on the heels of a prolonged period of stability, the more recent experience has offered painful evidence that banking crises do still occur.

policy prevents banking crises but requires, unrealistically, full knowledge of the shocks hitting the economy. By contrast, simple implementable rules that adjust capital requirements in response to a few observable macroeconomic indicators cannot always prevent crises. In this context, a static buffer turns out to be a sound, simple, and implementable policy.

Realistically, banking crises in our model have major consequences for consumption, business investment and household welfare. Bank capital requirements can curb the risk-taking incentives that lead to crises, and indeed changes to capital requirements continue to engage the policy and academic communities.[4] In our model, very high capital requirements force a bank to keep enough "skin in the game" to eliminate the excessive risk-taking incentives entirely. But capital requirements also reduce bank deposits, which provide liquidity services to households.

So, what are the optimal dynamic capital requirements? Triggering a risk-taking episode would lower household utility by a discrete amount. Accordingly, an all-knowing Ramsey planner—faced with aggregate and firm-specific shocks—would increase capital requirements just short of triggering a risk-taking episode. A less-informed regulator in the real world might be tempted to zip up to the regulator's estimate of the cliff's edge. But this well-meaning regulator could face a Wile E. Coyote moment, and may be better advised to exercise caution against banking crises, or do no harm.

We will explore this policy tradeoff both theoretically and quantitatively. We begin by showing how a Ramsey planner would respond to individual macroeconomic shocks, or a change in the volatility of returns in financial markets. We provide examples in which a Ramsey planner would raise capital requirements: (1) during a downturn caused by a productivity shock; (2) during an expansion caused by an investment-specific shock; or (3) during an increase in the volatility of financial market returns that has little effect on the business cycle. So, the all-knowing Ramsey planner will not necessarily set capital requirements in pro-, counter-, or a-cyclical manner. This is the basic reason why, as we shall see, simple counter-cycical rules for setting capital requirements do poorly in our model.

With many shocks hitting the economy simultaneously, the optimal Ramsey policy would require too much information to be implementable by a real-world regulator. Accordingly, we study the ability of simple and implementable policy rules to mimic the optimal policy.

More specifically, we use the simulated method of moments to calibrate our model's dynamic structure. This calibration allows us to calculate dynamic Ramsey capital requirements when the model economy is driven by a full constellation of shocks. And finally, we

[4]Following a pause in rule making during the Covid pandemic, many jurisdictions are moving to the finalization phase of the Basel III agreements. For the United States, Barr (2023) provides a preview of expected changes closely aligned with a recent notice of proposed rulemaking, Federal Register (2023). We review the related academic literature later in the introduction.

generate model data in that stochastic environment, and we regress the Ramsey capital requirements on variables that could be used for simple policy rules. The R-squared statistics from these regressions measure how well the simple rules mimic the optimal policy.

Of particular interest is the Basel III guidance for setting the countercyclical capital buffer (CCyB). According to this guidance, capital requirements should increase during periods of rapid credit expansion (or increases in the credit-to-GDP ratio), and they should be relaxed during a credit contraction.[5] This guidance—which we will call the "Basel rule"—sounds both sensible and implementable. And indeed, some statistical correlations would seem to support it. For example, Figure 1 suggests that the credit-to-GDP ratio (weakly) predicts GDP two years hence. U.S. data from the first quarter of 1980 through the fourth quarter of 2019, the blue circles in Figure 1, exhibit a negative correlation of about -0.1.[6]

Consistent with this statistical evidence, Figure 1 also shows that data from repeated simulations of our calibrated model imply a small negative correlation—repeated simulations of model data of the same size of the observed data yield a confidence set that includes the line based on the observed data.[7] In other words, our model is consistent with the empirical underpinnings that motivated the Basel III guidance on the CCyB.[8] Nonetheless, grounding those underpinnings in a theoretical framework leads to dramatically different policy prescriptions. The Basel III guidance does not come close to mimicking the optimal Ramsey policy in our model. An all-knowing Ramsey planner would tailor the policy response to the shocks hitting the economy and set capital requirements pro-, counter-, or a-cyclically depending on the shocks.

Actually, none of the simple rules we consider captures the Ramsey policy very well; that is, none has a high R-squared when the Ramsey policy is regressed on their variables. All of the simple rules let banks fall into the risk-taking trap from time to time, and we can calculate the frequency of these episodes.[9] We will see that a slightly elevated static capital requirement (or "buffer") largely avoids the Wile E. Coyote moments, and it does almost as well as the optimal rule when considering the level of deposits.

There are several strands of literature related to our work. The first is work in which

[5] The guidelines can be found in Basel Committee on Banking Supervision (2010).

[6] We will return to discuss the construction of Figure 1 in more detail in Section 7.1.

[7] For the model, the population correlation between the credit-to-GDP gap and the two-year-ahead GDP growth is -0.03, not far from the correlation of -0.1 for the observed sample from the first quarter of 1980 to the fourth quarter of 2019.

[8] Borio and Lowe (2002, 2004) provided the empirical underpinnings for the Basel III guidance. Their findings were bolstered by subsequent empirical work in Schularick and Taylor (2012), Jordà et al. (2017), Mian et al. (2017), and Mian et al. (2020).

[9] As a proof of concept, we show that complex rules, responding to a myriad of variables, fully capture the optimal rule, but still require unrealistic knowledge of the economy for their calibration, preventing their implementation.

endogenous regime shifts can lead to discontinous outcomes, financial crises, without relying on discontinuous distributions for exogenous shocks. See, for example, Mendoza (2010), He and Krishnamurthy (2012), and Brunnermeier and Sannikov (2014). Our approach also relies on endogenous regime shifts leading to banking crises, discontinuous events in which a large set of banks go bankrupt and deposit insurance intervenes to bail out depositors. Second, there is a literature on credit booms and busts, including Boissay et al. (2016) and Bordalo et al. (2018). We differ mainly by not (explicitly) modeling banking panics and by integrating the analysis within a reasonably conventional quantitative macroeconomic framework with a clear role for capital regulation of banks.

Several influential contributions to the literature emphasize risks arising from high leverage and the expansion of bank credit. Davydiuk (2017) and Malherbe (2020) are examples of this. Our work offers a complementary perspective that emphasizes the composition of bank credit, rather than its expansion. Gomes et al. (2023) develop a model that shares our emphasis on risks arising from changes in the composition of bank credit, rather than the expansion of credit. They question the premise that high leverage causes banking crises and that curbing credit growth can prevent crises. In their model, a time-varying likelihood of an exogenous economic crisis causes both higher leverage and the subsequent economic decline, making policies that respond to credit growth ineffective.

The papers by Begenau (2020), Begenau and Landvoigt (2022), Collard et al. (2017), Martinez-Miera and Suarez (2014), and Panscott and Robatto (2022) share the risk-shifting framework in our model. Begenau (2020) and Panscott and Robatto (2022)—arguably the closest analyses to our own—focus on the optimal level of static capital requirements in the steady state. Our focus is on cyclical capital requirements, and their comparison with static buffers. Begenau and Landvoigt (2022) center on the interactions between regulated and unregulated banks. Martinez-Miera and Suarez (2014) develop a model with systemic risk in the form of a binary shock, which simplifies the theoretical derivations and numerical solution. By contrast, our model has a richer stochastic structure, important for our assessment of simple rules, something they do not attempt. Collard et al. (2017) concentrate on interactions of optimal monetary and prudential policies, in a setting that keeps bank failures off the equilibrium path. We abstract from monetary policy, but we allow for business-cycle fluctuations and risk taking on the equilibrium path.

Finally, other contributions show how leverage can increase financial fragility and the risk of bank runs. Examples include Angeloni and Faia (2013), Gertler and Kiyotaki (2015), Faria-e-Castro (2019), and Gertler et al. (2020). We make our formal analysis stark by setting aside bank runs, but of course we recognize the possibility of bank runs in reality.

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. A series of appendices provide more detailed derivations of some of our results, and Appendix H performs important sensitivity analyses; Section 8 summarizes the results of Appendix H. And Section 9 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, non-financial 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 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^s_{t+1}, to produce

$$y^s_{t+1} = A_{t+1}(k^s_{t+1})^{\alpha}(h^s_{t+1})^{1-\alpha}, \tag{2}$$

where A_{t+1} is an aggregate shock to total factor productivity (TFP). When a safe firm takes the loan in period t, it knows that the firm will hire the optimal h^s_{t+1} next period. So, the safe firm chooses $l^{f,s}_t$ and k^s_{t+1}in period t, and then h^s_{t+1} in period $t+1$, to

$$\max_{l^{f,s}_t, k^s_{t+1}} E_t \left\{ \max_{h^s_{t+1}} \left[y^s_{t+1} + (1-\delta)Q_{t+1}k^s_{t+1} - W_{t+1}h^s_{t+1} - R^s_{t+1}l^{f,s}_t \right] \right\}, \tag{3}$$

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

$$E_t R^s_{t+1} = \alpha E_t \left\{ \frac{A_{t+1}}{Q_t} \left(\frac{h^s_{t+1}}{k^s_{t+1}} \right)^{1-\alpha} + (1-\delta)\frac{Q_{t+1}}{Q_t} \right\}, \tag{4}$$

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 $y^r_{t+1} = A_{t+1}\left(k^r_{t+1}\right)^{\alpha}\left(h^r_{t+1}\right)^{1-\alpha} + \varepsilon_{t+1}k^r_{t+1}$, where ε_{t+1} is an idiosyncratic shock that follows a Normal distribution G with a negative mean, $-\ \xi$, and standard deviation τ:[10]

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

The risky firm chooses $l^{f,r}_t$ and k^r_{t+1}, and then h^r_{t+1}, to

$$\max_{l^{f,r}_t, k^r_{t+1}} E_t \left\{ \max_{h^r_{t+1}} \left[y^r_{t+1} + (1-\delta)Q_{t+1}k^r_{t+1} - W_{t+1}h^r_{t+1} - R^r_{t+1}l^{f,r}_t \right] \right\}, \tag{6}$$

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

$$E_t R^r_{t+1} = E_t R^s_{t+1} - \frac{\xi}{Q_t}. \tag{7}$$

[10] $\exp(x) = e^x$ is the exponential function and $\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}}\int_{-x}^{x} \exp\left(-v^2\right) dv = \frac{2}{\sqrt{\pi}}\int_0^x \exp\left(-v^2\right) dv$.

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 given by $(1-\alpha)A_{t+1}(k^i_{t+1}/h^i_{t+1})^\alpha$ 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.

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

Appendix B.3 provides the details on aggregation across firms; 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

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

where capital letters represent aggregate values.

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 a shock to investment-specific technology (ISP), and ϕ is a measure of the severity of investment adjustment costs.[11] 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

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

[11]We include investment adjustment costs, and later habits in consumption, to make our model fit the data better. But they are not an integral part of the logic behind capital requirements.

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

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

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; β is the discount factor; $0 < \kappa < 1$ is the habit persistence parameter, $\varsigma_c > 0$ captures the intertemporal elasticity of substitution, $\varsigma_0 > 0$ is the utility weight on deposits, and $\varsigma_d > 0$ is the inverse elasticity of household 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. For simplicity, we assume that households supply labor inelastically, and we have normalized the supply of labor to be one.[12] 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:

$$C: \quad (C_t - \kappa C_{t-1})^{-\varsigma_c} - \beta \kappa E_t (C_{t+1} - \kappa C_t)^{-\varsigma_c} - \lambda_{ct} = 0, \tag{16}$$

$$D: \quad \varsigma_0 D_t^{-\varsigma_d} - \lambda_{ct} + E_t \beta \lambda_{ct+1} R_t^d = 0, \tag{17}$$

$$E^s: \quad -\lambda_{ct} + E_t \beta \lambda_{ct+1} R_{t+1}^{e,s} + \zeta_t^s = 0, \tag{18}$$

$$E^r: \quad -\lambda_{ct} + E_t \beta \lambda_{ct+1} R_{t+1}^{e,r} + \zeta_t^r = 0, \tag{19}$$

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

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.

2.3 Banks

Banks are at the heart of our model. First, we set the stage by describing their incentives to take excessive risk. Then, 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 $E_t R_{t+1}^r < E_t R_{t+1}^s$. 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 are born each period, and they live for two periods. In the first period, a bank issues equity, e_t, and deposits, d_t, to households, and uses the proceeds to make loans, l_t, 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, a bank incurs a cost of originating and monitoring its loans, fl_t, where f is the cost per loan. 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.[13] Since $R^r_{t+1} = R^s_{t+1} + \frac{\varepsilon_{t+1}}{Q_t}$, the ex-post return on the portfolio will be $R^s_{t+1} + \sigma_t \frac{\varepsilon_{t+1}}{Q_t}$.

The bank's net worth in period $t+1$ consists of its earnings on the loan portfolio net of the interest payments on its deposits:

$$nw_{t+1} \equiv \left(R^s_{t+1} - f + \sigma_t \frac{\varepsilon_{t+1}}{Q_t} \right) l_t - R^d_t d_t. \tag{20}$$

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 below which the bank's net worth is negative; that is, $\left(R^s_{t+1} - f + \sigma_t \frac{\varepsilon^*_{t+1}}{Q_t} \right) l_t - R^d_t d_t = 0$. 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:

$$\max_{l_t, d_t, e_t, \sigma_t} E_t \left\{ \psi_{t.t+1} \left[\int_{\varepsilon^*_{t+1}}^{\infty} nw_{t+1} \, \mathrm{d}G(\varepsilon_{t+1}) \right] \right\} - e_t, \tag{21}$$

subject to

$$\begin{aligned} l_t &= e_t + d_t, \\ e_t &\geq \gamma_t l_t, \\ l_t &\geq 0, \\ \underline{\sigma} &\leq \sigma_t \leq \bar{\sigma}, \end{aligned} \tag{22}$$

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

[13] 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.

they so choose.[14]

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 Risk

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

$$\varOmega(\sigma_t;\, l_t,\, d_t,\, e_t) = E_t\left[\psi_{t,t+1} l_t\left(\omega_1 + \omega_2\right)\right], \tag{23}$$

where

$$\omega_1 \equiv \left(R^s_{t+1} - f - R^d_t\left(1-\gamma_t\right) - \frac{\xi\sigma_t}{Q_t}\right)\left(1 - G(\varepsilon^*_{t+1})\right), \tag{24}$$

$$\omega_2 \equiv \left(\frac{\sigma_t}{Q_t}\right)\frac{\tau}{\sqrt{2\pi}}\exp\left(-\left(\frac{\varepsilon^*_{t+1}+\xi}{\tau\sqrt{2}}\right)^2\right), \tag{25}$$

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, we show the tax necessary to support the insurance scheme is

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

$$T_t = \frac{\sigma_{t-1}L_{t-1}}{Q_{t-1}} \frac{\tau}{\sqrt{2\pi}} \exp\left(- \left(\frac{\left(R^d_{t-1}(1-\gamma_{t-1})+f-R^s_t\right)Q_{t-1}+\xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau} \right)^2 \right) - \tag{26}$$
$$\frac{1}{2}L_{t-1}\left(R^s_t - f - \frac{\sigma_{t-1}\xi}{Q_{t-1}} - R^d_{t-1}\left(1-\gamma_{t-1}\right)\right)\left[1 + \operatorname{erf}\left(\frac{\left(R^d_{t-1}(1-\gamma_{t-1})+f-R^s_t\right)Q_{t-1}+\xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)\right],$$

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 the 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 A.2.

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.

Note also that a risky bank seeks to maximize its exposure to the idiosyncratic shock ε_t. Limited liability incentivizes banks to "fail big." So, a risky bank would not want to diversify its loan portfolio by lending to more than one risky firm.[15]

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 conditions linking our bank-level and firm-level variables representing loans are

$$Q_t K_{t+1}^s = (1 - \underline{\sigma})(1 - \mu_t) l_t^s + (1 - \bar{\sigma}) \mu_t l_t^r. \tag{27}$$

Similarly,

$$Q_t K_{t+1}^r = \underline{\sigma}(1 - \mu_t) l_t^s + \bar{\sigma} \mu_t l_t^r. \tag{28}$$

The aggregate bank loans are linked to the individual bank loans by: $L_t^r = \mu_t l_t^r$ and $L_t^s = (1 - \mu_t) l_t^s$. Therefore, we can describe the latter two equations by using aggregate loans

$$\begin{aligned} Q_t K_{t+1}^s &= (1 - \underline{\sigma}) L_t^s + (1 - \bar{\sigma}) L_t^r, & (29) \\ Q_t K_{t+1}^r &= \underline{\sigma} L_t^s + \bar{\sigma} L_t^r. & (30) \end{aligned}$$

The equity positions taken by households, in turn, determine the equity positions of individual banks: $E_t^r = \mu_t e_t^r$ and $E_t^s = (1 - \mu_t) e_t^s$. The returns on the equity positions taken

[15] In reality, bank regulators would not allow a bank to lend to a single firm. But our result really says that risky banks seek exposure to a single idiosyncratic shock ε_t. To circumvent regulation, for example, a bank may hold a seemingly diversified portfolio of MBS with all the loans exposed to the risk of a decrease in house prices. These incentives seem relevant for the literature on Securitization surveyed by Gorton and Metrick (2013).

by households at date t are linked to the dividends paid by banks at date $t+1$. We have:

$$E_t^r R_{t+1}^{e,r} = (\omega_1^r + \omega_2^r)\, L_t^r, \tag{31}$$

$$E_t^s R_{t+1}^{e,s} = (\omega_1^s + \omega_2^s)\, L_t^s, \tag{32}$$

where we use the fact that max $\left[nw_{t+1}^r, 0\right]$ is linear in loans; ω_1 and ω_2 were defined in equations (24) and (25). Deposits held by households are issued by (safe and risky) banks: $D_t = D_t^s + D_t^r$ where $D_t^s = L_t^s - E_t^s$ and $D_t^r = L_t^r - E_t^r$.

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

$$H_t^s + H_t^r = 1, \tag{33}$$

$$K_t^s + K_t^r = K_t, \tag{34}$$

and

$$Y_t^s + Y_t^r = C_t + I_t^g. \tag{35}$$

3 Calibration of Parameters and Optimal 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. Our setup for investment adjustment costs mimics the one used by Altig et al. (2011). We pick a value of ϕ consistent with the broad range from their analysis and related literature.

Two parameters that govern the attractiveness of excessively risky loans are specific to our model: τ is the standard deviation of the risky firm's idiosyncratic shock, and ξ is the average penalty for financing risky projects.[16] A higher value for τ makes risky loans more attractive (by further exploiting limited liability) and a higher value for ξ makes risky loans less attractive.

To choose these parameters in an empirically relevant way, we rely on a definition of excessive risk that was agreed upon by three regulators of U.S. depository institutions—the Office of the Comptroller of the Currency, the Board of Governors of the Federal Reserve,

[16] More precisely, $-\xi$ is the expected return on a risky loan.

and the Federal Deposit Insurance Corporation. In March of 2013, these regulators issued guidance on leveraged lending with the aim of ensuring that financial institutions did not "unnecessarily heighten risks by originating poorly underwritten loans."[17] This guidance established a bright line that loans to firms with a debt-to-EBITDA ratio of 6 or above would raise supervisory concerns.[18]

We choose τ to make the variance of returns on a risky project match the variance of returns from lending to a firm with a debt-to-EBITDA ratio of 6. We focus on variances conditional on starting from the non-stochastic steady state. Given τ, we choose the value of ξ to make $\gamma = 0.10$ be the steady-state capital requirement that is just high enough to prevent lending to risky firms. We note that 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.[19]

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).

Finally, the parameters ς_0 and ς_d appear in the household's utility of deposits. The values of these parameters are potentially important to our inquiry, since the fundamental tradeoff for our Ramsey planner is between the utility of deposits and the disutility of excessive risk taking. ς_0 measures the importance of the utility of deposits relative to the utility of consumption. We choose the value of ς_0 to make the steady-state interest rate on bank deposits equal to 0.86% per quarter, a value we borrow from an estimate in Begenau (2020); in particular, we set $\varsigma_0 = 0.25$. The willingness of households to vary their supply of deposits as consumption or deposit rates move is governed by the parameter ς_d; the lower is this parameter value, the more willing are households to adjust deposits to cushion fluctuations in consumption. We set $\varsigma_d = 1.1$, a numerical approximation of the log case. We discuss the sensitivity of our results to this parameter in Section 8.

4 Numerical Methods

Occasionally binding non-negativity constraints on bank loans complicate the solution of our model. To address these complications, we rely on the OccBin toolkit developed by Guerrieri and Iacoviello (2015); they also provide an extensive discussion of the accuracy of their solution. Their algorithm can be applied to models with a large number of state variables, such as ours.

[17] The inter agency guidance on leveraged lending can be found here.

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

[19] In Section 8, we explain why we do not use our model to directly calculate an optimal steady-state capital requirement and offer extensive sensitivity analysis.

So why did we complicate matters by imposing non-negativity 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. The last possible scenario is when safe and risky loans are equally priced, so the expected returns on safe and risky loans are the same, resulting in a mixed equilibrium in which $0 < \mu_t < 1$ (as described in Section 2.5.2).

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 $\{\gamma_t^A\}_{t=0}^{\infty}$ in which $\gamma_t^A = \gamma_t^*$ for $t \neq t_k$ and $\gamma_t^A = \gamma_t^* + \Delta$ for $t = t_k$ and some $\Delta \neq 0$. When $\Delta > 0$, $\{\gamma_t^A\}_{t=0}^{\infty}$ 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 $\{\gamma_t^A\}_{t=0}^{\infty}$. 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 to the lowest possible levels necessary to prevent risk-taking $\{\gamma_t^*\}_{t=0}^{\infty}$ otherwise.[20] 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.

6 Optimal Dynamic Capital Requirements

In this section, we show how a Ramsey planner would set capital requirement ratios, γ_t, in response to three shocks that have different cyclical implications for GDP and credit. Specifically, we show that optimal capital requirements can increase in a recession or a boom, or may adjust to prevent a banking crisis in response to shocks that leave little inprint on GDP. We also show that the optimal capital requirements exhibit different patterns of correlation with the credit-to-GDP ratio depending on the source of shocks. Regardless of these cyclical differences, in all cases, the Ramsey planner will set capital requirements to prevent a banking crisis.

We take two steps in preparation for our discussion here. First, we ask what might trigger a risk-taking episode that leads to a banking crisis. Then, to isolate the effects

[20] For each of our checks, we recompute the path of the minimum capital requirements that prevent risk as this path also depends on the evolution of the endogenous variables in the regime with excessive risk taking.

of moving capital requirements, we consider how exogenous changes in these requirements would affect financing decisions and real allocations.Finally, with this preliminaries out of the way, we consider shocks to aggregate productivity, investment productivity, and volatility of the idiosyncratic productivity for risky firms.[21] We also contrast the repercussions of each shock under the optimal Ramsey policy this those under a setup that leaves capital requirements unchanged.[22] We leave a systematic assessment of simple and implementable policies, including a static buffer, for Section 8.

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 (23) specifies $\mathit{\Omega}(\sigma_t;\, l_t,\, d_t,\, e_t)$ for all values of σ_t, where it will be recalled that

$$\varepsilon_{t+1}^{*} = -\frac{Q_t}{\sigma_t}\left[R_{t+1}^{s} - f - R_t^d\left(1-\gamma_t\right)\right] \tag{36}$$

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, (21), 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 suppose that $\underline{\sigma} = 0$ and $\bar{\sigma} = 1$ in this subsection. With these simplifications, (23) implies

$$\Omega_t^s = E_t\left[\psi_{t,t+1} l_t^s\left(R_{t+1}^s - f - R_t^d\left(1-\gamma_t\right)\right)\right] \quad \text{and} \tag{37}$$

$$\begin{aligned}\Omega_t^r = E_t\Bigg[\psi_{t,t+1} l_t^r\Bigg(&\left(R_{t+1}^s - f - R_t^d\left(1-\gamma_t\right) - \frac{\xi}{Q_t}\right)\left(1 - G(\varepsilon_{t+1}^{*})\right) + \\ &\frac{\tau}{Q_t\sqrt{2\pi}}\exp\left(-\left(\frac{\varepsilon_{t+1}^{*}+\xi}{\tau\sqrt{2}}\right)^2\right)\Bigg)\Bigg]\end{aligned} \tag{38}$$

where it will be recalled that

[21] All the shocks will follow an auto-regressive process of order 1 with persistence parameters spelled out with the description of effects of each shock. The rest of the parameter settings are given in Table 1, except that here we set $\varphi = 100$ and $\kappa = 0$.

[22] 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 choice simplifies our numerical solution methods by avoiding plunging the economy into a banking crisis even for tiny misallagnments in capital requirements.

$$R^s_{t+1} = \alpha \left\{ \frac{A_{t+1}}{Q_t} \left(\frac{H^s_{t+1}}{K^s_{t+1}} \right)^{1-\alpha} + (1-\delta) \frac{Q_{t+1}}{Q_t} \right\}. \tag{39}$$

What might turn $\Omega^r_t - \Omega^s_t$ positive, triggering a risk-taking episode? The obvious culprit is the interest rate spread $R^s_{t+1} - f - R^d_t (1-\gamma_t)$. 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: (36) implies that ε^*_{t+1} will rise. The presence of ε^*_{t+1} (instead of $-\infty$) in the bank's expected profits, (21), 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.[23] 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 (38) 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, $-\xi$; 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 its 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 not in

[23] It is hard to see these results in (38) 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.

a risk taking episode, 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 spread in asset returns—whose evolution is closely related to $\Omega_t^r - \Omega_t^s$—that we can track:

$$S_t \equiv E_t \left[R_{t+1}^{e,r} - R_{t+1}^{e,s} \right]. \tag{40}$$

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 of 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.[24]

6.2 Capital Requirement Shocks

The next two sections illustrate the transmission mechanism for changes to capital requirements. 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 2 shows the effects of a one percentage point increase in the capital requirement, 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.

[24] 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, $S_t \equiv E_t \left[R_{t+1}^{e,r} - R_{t+1}^{e,s} \right] = \frac{\Omega_t^r}{E_t^r} - \frac{\Omega_t^s}{E_t^s}$. 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 2. 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.[25] 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 2.[26]

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 3 show the response to a one percentage point 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.[27] 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

[25]Begenau (2020) also finds that an increase in capital requirements can reduce the cost of bank funding and increase lending.

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

[27]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 (39); 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 3, repeat the responses shown in Figure 2; 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 A Contractionary TFP Shock

TFP shocks have played a major role in RBC modeling. Figure 4 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, leading to the failure of a subset of banks; 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 4. 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 it responds 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 of ξ, the expected loss on risky loans; ξ is a measure of the economic inefficiency in our model.

To set the stage for the analysis of simple rules, including the Basel III CCyB prescriptions, is is useful to focus on the response of the credit-to-GDP ratio. Figure 4 shows that this ratio falls on impact, then rises overshooting its steady state before gradually falling back down. This non-monotonic response is in contrast with the smooth response of the Ramsey optimal capital requirements and presages that a simple rule for capital requirements responding to the credit-to-GDP ratio (or its gap from trend) would need to include not just the current value but also lags (and possibly a large number of lags) of this ratio to approximate the optimal response.

Takeaways:

The optimal Ramsey policy avoids a banking crisis in the face of an economic contraction caused by a negative TFP shock with an increase in capital requirements. Following a TFP shock, there is no simple pattern linking the responses of the credit-to-GDP ratio and of the Ramsey optimal capital requirements.

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 5 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 an associated banking crisis 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.

In response to this shock, just as for the TFP shock of Section 6.3, the credit-to-GDP ratio does not move monotonically. On impact it drops, then it moves up overshooting its steady state. In this case, the overshooting is more persistent than for the TFP shock, so there is a longer period in which the credi-to-GDP ratio is rising while the capital requirements are moving back down toward the steady state. Once again, the rich dynamics do not point to a simple, time-invariant correlation between the Ramsey optimal capital requirements and the credit-to-GDP ratio (or its gap relative to trend). Moreover, the complex correlation pattern also varies depending on the shock affecting the model economy.

Takeaways:

The optimal Ramsey policy avoids a banking crisis in the face of an economic boom caused by an investment technology shock with an increase in capital requirements. There is no time-invariant contemporaneous correlation between the Ramsey optimal capital requirements and the credit-to-GDP ratio and the correlation pattern also varies depending on which shock or combination of shocks is affecting the economy.

6.5 A Shock that Increases the Volatility of Risky Returns

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 6 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 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.

Finally, returning to the response of the credit-to-GDP ratio, the optimal Ramsey policy for capital requirements leaves it essentially unchanged in response to the volatility increase.
Takeaways:

The optimal Ramsey policy can neutralize the effects of shocks that affect the desirability of risky projects by raising capital requirements leaving little to no imprint on GDP and the credit-to-GDP ratio. This is further evidence that the correlation pattern between optimal capital requirements and the credit-to-GDP ratio is shock dependent.

7 Moment Matching and Simple and Implementable Rules for Capital Requirements

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. In our baseline calibration, we use the volatility shock and the two macroeconomic shocks—TFP and ISP. 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 the 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 this calibration, 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 2,000 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.[28] 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. The Ramsey policy varies capital requirements to avoid excessive risk-taking episodes, otherwise having little impact on the macroeconomy.

7.1 Matching Moments, Shock Processes and Variance Decompositions

Table 4 shows that our calibration is very good; model moments are close to data moments. Tables 2 and 3 show the calibrated shock processes and the variance decompositions. It may be interesting to note that the persistence parameter for the TFP shock is 0.79, which is somewhat lower than what is normally assumed in the RBC literature.[29] Finally, the parameters for consumption habits and investment adjustment costs that minimize the distance function are 0.93 and 0.06 respectively.

In our calibration, all of the shocks are persistent. In the variance decompositions, the TFP shock explains all of the variations in GDP and investment, while the volatility shock explains the variation in the Ramsey policy settings.

[28] Why the Ramsey policy? We are calibrating to data from the pre-crisis period; the period between S&L crises and 2008 did not have great bank failures. Either capital requirements were high enough, or shocks were small enough, to avoid risk taking. In the context of our model, the Ramsey policy captures this.

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

Given our interest in assessing simple rules for capital requirements that respond to the ratio of credit to GDP, we find it important to assess how well we can capture key data relationships that underpin the recommendation to consider this type of rule in the Basel III Accords. Accordingly, returning to Figure 1, the figure shows that there is a predictive relationship between the ratio of non-financial credit to GDP and GDP two years ahead using data for the United States. The circles in Figure 1 denote observations from the first quarter of 1980 through the fourth quarter of 2019, pointing to a tenuous negative correlation, about -0.1. The regression line in the figure confirms that periods in which the credit-to-GDP ratio is above trend systematically presage periods in which GDP will be below trend two years into the future.[30]

To assess whether or not our model is consistent with the negative correlation from U.S. data, we draw from the model 1,000 simulated samples of the same length as the observed data. For each sample we recompute the regression line shown in Figure 1 and the correlation between the detrended ratio of credit to GDP and GDP two years ahead, which allows us to size a 90 percent confidence interval. Both are encompassed in their respective confidence intervals. Specifically, the interval for the correlation runs approximately from -0.2 to 0.2, thus including the -0.1 correlation based on observed data. We conclude that, even though this is not a moment directly targeted in our calibration based on the simulated method of moments, the model is consistent with the mild predictive relationship that can be evinced from U.S. data.[31]

7.2 Simple and Implementable Rules for Capital Requirements

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).

[30] We detrended the credit-to-GDP ratio with a Hodrick-Prescott filter using a coefficient of 400,000, implying a trend that is almost linear in line with the prescriptions of the Basel III guidance. This trend is consistent with the recommendations in Borio and Lowe (2002) and Basel III guidance, Basel Committee on Bank Supervision (2010).

[31] Regression results analogous to the ones presented here obtain for various sensitivity exercises, including: a longer sample starting in 1947, using detrended non-financial business credit without dividing it by GDP, and lengthening or shortening the lead length for detrended real GDP.

Then, we use a variety of easily interpretable measures to rank the alternative rules. The first, and perhaps the most obvious, measure is the R-squared of the regression; the higher the R-squared, 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.[32]

In the assessment of simple rules presented in the next section, we impose a small safety corridor next to the cliff of excessive risk taking in the form of a static buffer—without this buffer the simple rules we study would plunge the economy into banking crises even more often. We consider static buffers with alternative sizes, 10, 50, and 100 basis points. In other words, we push up the steady-state capital requirement from 10 percent to 10.1, 10.5, and 11 percent, respectively.

7.2.1 Simple Rules

We start this exploration by focusing on the Basel rule and showing that it does very poorly. The Basel III prescription is to tighten or relax capital requirements in line with changes in the credit-to-GDP ratio.[33] As shown in Table 5, the R-squared for this rule is only 0.016. Moreover, the prevalence of risk-taking quarters is high with the static buffer at 10 basis points. The reason for the poor performance is in line with the analysis presented in Section 6 which highlighted two types of problems: First, there is a complex correlation pattern between optimal capital requirements and the credit-to-GDP ratio, which varies over time in response to shocks; and second, this correlation pattern is shock dependent. We confirmed that, if the economy were only affected by one type of shocks, say TFP shocks, a rule responding to the credit-to-GDP ratio and its first and second lags would nearly reproduce the optimal Ramsey rule (the R-squared would be 0.99). Alas, the same kind of rule would require many more lags if the economy were only buffeted by ISP shocks.

[32] Focusing on multiple easily interpretable performance measures rather than one summary welfare metric facilitates the interpretation of our results.

[33] The Basel III guidance on the CCyB recommends detrending the credit-to-GDP ratio with a Hodrick-Prescott filter with a parameter set to 400,000, implying a trend very close to linear. Since our model is stationary, there is little difference between the ratio and this gap. For simplicity, we focus on a rule in terms of the ratio.

Moreover, in a more realistic setting with multiple shocks, including volatility shocks, the coefficients in the rule would have to vary depending on the combination of shocks affecting the economy in any one period, or the rule would have to depend on the size of the shocks, their lags, and innovations, to match the optimal response of the Ramsey policy, making the rule neither simple nor implementable.

Note also that the regression coefficient is small in magnitude and negative in sign, not in line with the Basel III recommendations. We can offer a little more intuition for this finding by computing unconditional correlations (in population) between capital requirements and the credit-to-GDP ratio conditional on different shock sources. At 0.67, this correlation is indeed positive and sizable when only turning on TFP shocks. Switching to ISP shocks, the same correlation takes on a value of -0.68, almost equal in magnitude but opposite in sign than for the case of TFP shocks. And when only conditioning on volatility shocks, the covariance is 0.[34] One can easily see that a simple rule (one that cannot vary depending on the mix of underlying shocks) will comingle the responses to different shocks in a way that will result in a muted reaction—the small regression coefficient in Table 5. This muted reaction will end up being suboptimal regardless of the shock, as shown in the table.

Table 5 also shows that a simple rule for capital requirements that imposes a positive coefficient on the credit-to-GDP gap fares even worse, as expected. Raising the static buffer to 100 basis points improves the performance of the Basel rule. But the higher steady-state capital buffer is doing all the work, and the performance is still far from that of the optimal Ramsey policy.

Beyond showing the performance of the Basel III rule and closely related variants, Table 5 covers the performance of a broader set of simple and implementable rules. How much worse do other implementable rules do?

Among the simplest rules that only respond to one variable, the rule that reacts to the expected banking spread (the spread between the expected interest rate on safe loans and the interest rate on deposits) is the closest match to the optimal Ramsey rule. The idea of moving capital requirements with the expected banking spread is not without merit. In line with the discussion on the triggers of risk-taking episodes in Section 6, a higher expected banking spread makes risky projects less attractive by decreasing the value of the shield of limited liability.[35] However, this simple rule is still very far from meeting the Ramsey performance standards—the number of risk-taking quarters per 100 years is very high when

[34] To compute the population correlations conditional on different shock sources, we simulate three samples of 10,000 observations drawing innovations for only one of the three shock sources in each sample. We compute the credit-to-GDP gap using a Hodrick-Prescott filter with a coefficient of 400,000.

[35] We verified that the expected banking spread goes up for all three shocks—TFP, ISP, and volatility—considered in Section 6.

the steady-state capital requirement and the average level of deposits is very low. There are dramatic improvements if the rule is used in combination with the higher static buffers considered in Table 5, but the buffers are doing all the work.

It may seem counterintuitive that pushing up the capital buffer may increase the level of deposits; after all, higher capital requirements shift bank financing away from deposits, which are valued by households. However, consider that both investment and loans shrink in a banking crisis. Limiting the number of risk-taking episodes increases the average amount of credit that is extended, which can raise the level of deposits even when deposits account for a lower fraction of bank funding. In that vein, one might surmise that rules for capital requirements responding to measures of aggregate activity might work well. Alas, Table 5 shows that a rule responding to GDP fares no better than the actual Basel rule. The R-squared is virtually zero; so, the simple rule is not tracking the Ramsey policy, and the performance measures are poor.

To proceed systematically, we also fit rules to each of the state variables from the model solution and to all possible combinations of two of those variables. There is no success in this endeavor, and we do not show these additional results in Table 5.

The remaining rules are not implementable because of their informational requirements, but we offer them as a proof of concept. If we could observe the shocks and their innovations, we ought to be able to shift the cyclical properties of the rule to match the optimal rule. As conjectured, when armed with all this information, the R-squared statistics are 1.0, and the rules come close to matching the performance of the Ramsey policy—the only reason for a slightly lower average level of deposits is the imposition of static buffers.

To conclude, simple instrument rules do not come close to matching the performance of the optimal rule. Complicated rules do, but those rules are not implementable. The addition of a static buffer seems to improve the performance of simple rules. So, we study static buffers next.

7.2.2 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 toolkit. Table 6 bears that out. Here, there are no rules, just static capital buffers. The last row reports 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 11.5 percent, the results are almost as good as those achieved by the Ramsey planner and much better than the performance measures for any of the implementable rules in Table 5. These results suggest that the regulator need not bother with dynamic capital requirements.

8 Sensitivity Analysis

In this section, we discuss the sensitivity of our results to various parameter settings, and to alternative calibrations of the shock processes that drive the model's economy. The details are pushed to Appendix H.

8.1 The Optimal Steady-State Capital Requirement

As noted in the calibration section, Section 3, there are two parameters that are specific to our model: τ is the standard deviation of the risky firm's idiosyncratic shock, and ξ is the average penalty for financing risky projects.[36] We chose τ to fit the data, and then we treated ξ as a free parameter. We chose ξ to pin down an empirically plausible optimal steady-state capital requirement; that is, we set ξ to make $\gamma = 0.10$.

Why did we not try to choose ξ empirically, and then calculate an optimal steady-state capital requirement directly? We show in Appendix H.1 that small variations in τ or ξ would support a wide range of steady-state capital requirements. For example, steady-state capital requirements vary from about 5% to 15% when ξ is chosen from a very narrow range. τ and ξ cannot be credibly estimated with that kind of precision, suggesting that our model is not suitable for a serious attempt to pin down the optimal steady-state value.

8.2 Volatility of Optimal Dynamic Capital Requirements

How much do optimal dynamic capital requirements have to be adjusted? In Appendix H.2, we explore the relative volatility of the optimal dynamic responses implied by three parameters: τ, ξ, and ς_d (the inverse of the interest rate elasticity of the household's supply of bank deposits). As an example, we focus on the dynamic response to a TFP shock. We show that an increase in τ, or a decrease in ξ, require a larger adjustment in the optimal capital requirements but no change in their cyclical properties. These results may not be too surprising, since these parameter changes increase the attractiveness of risky loans.

In line with related papers, we choose ς_d to imply an interest elasticity of deposit supply close to 1. Appendix H.3 shows that our results are not sensitive to lowering the elasticity in the range between 0.15 and 1. As we raise the elasticity towards infinity, however, the risk-taking incentives grow, and the optimal capital requirements becomes more volatile. We don't think this is an empirically relevant result.

[36]More specifically, $-\xi$ is the expected loss on a risky loan.

8.3 An Alternative Calibration of Shock Processes

In Section 7.1, we calibrated shock processes to make our model's moments match those found in the U.S. data. We then used the calibrated model to evaluate implementable policy rules for dynamic capital requirements. Under our benchmark calibration, static capital buffers dominated these policy rules.

The first step in the procedure was to decide which shocks drive the macroeconomy. In our baseline calibration, we used the volatility shock and the two macroeconomic shocks—TFP and ISP. The resulting model was very good at matching moments in the data.

The choice of shocks is, however, not innocent. In Appendix H.4, we consider an alternative calibration that only uses the two macroeconomic shocks—TFP and ISP. Then we show that this calibration is just as good as our benchmark calibration in terms of matching moments. In this model, the ISP shock is an important driver of our dynamic capital requirements, and indeed a simple rule based on the price of investment, Q_t, tracks the Ramsey policy fairly well. However, the Basel rule still performs badly, and a static buffer looks even more effective.

9 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. We also show that depending on the underlying shocks, the correlation between the Ramsey optimal capital requirements and the credit-to-GDP gap can be large in magnitude and positive or negative in sign, or there can be no covariance.

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 reliable indicators. To this end, we employed a 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 a rule following the Basel III guidance on the setting of countercyclical capital buffers—may sound sensibly grounded in empirical regularities but turn out to do more harm than good in our model. We find it useful to ground the policy analysis on theoretical underpinnings consistent with those very same empirical regularities. These underpinnings allow us to show that fine tuning capital requirements can be 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

Note: See Section 3 for the calibration strategy.

Table 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	

Note: This table shows the parameter values of the shock processes estimated by a SMM procedure. We use the volatility shock and the two macroeconomic shocks—TFP and ISP (investment specific). The distance function reports the distance between the model and data moments weighted by the matrix that minimizes the variance of the estimates (optimal matrix).

Table 3: 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

Note: This table shows the variance decompositions of GDP, gross investment, investment price, capital requirement, and the credit-to-GDP ratio.

Table 4: 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

Note: This table compares data and model moments—variances, correlations, and autocorrelations—in our calibration (using a SMM estimation procedure).

Table 5: Simple Rules

Simple Rule	R Squared	Regression slope coefficient	Static buffer = 10 basis points: Quarters with excessive risk-taking (per 100 years)	Static buffer = 10 basis points: Average deposit under simple rule	Static buffer = 50 basis points: Quarters with excessive risk-taking (per 100 years)	Static buffer = 50 basis points: Average deposit under simple rule	Static buffer = 100 basis points: Quarters with excessive risk-taking (per 100 years)	Static buffer = 100 basis points: Average deposit under simple rule
1. Investment price	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-to-GDP gap	0.016	-0.005	208.4	7.777	76.8	13.027	7.2	15.788
5. Credit-to-GDP gap with positive coefficient		0.005	Convergence problems		83.2	12.780	6.8	15.805
6. All shock processes, innovations, expected safe return and deposit rate	1.000	Too many to show	0	16.223	0.0	16.151	0	16.061
7. All shock processes, innovations, and lagged capital requirement	1.000	Too many to show	0	16.223	0.0	16.151	0	16.061

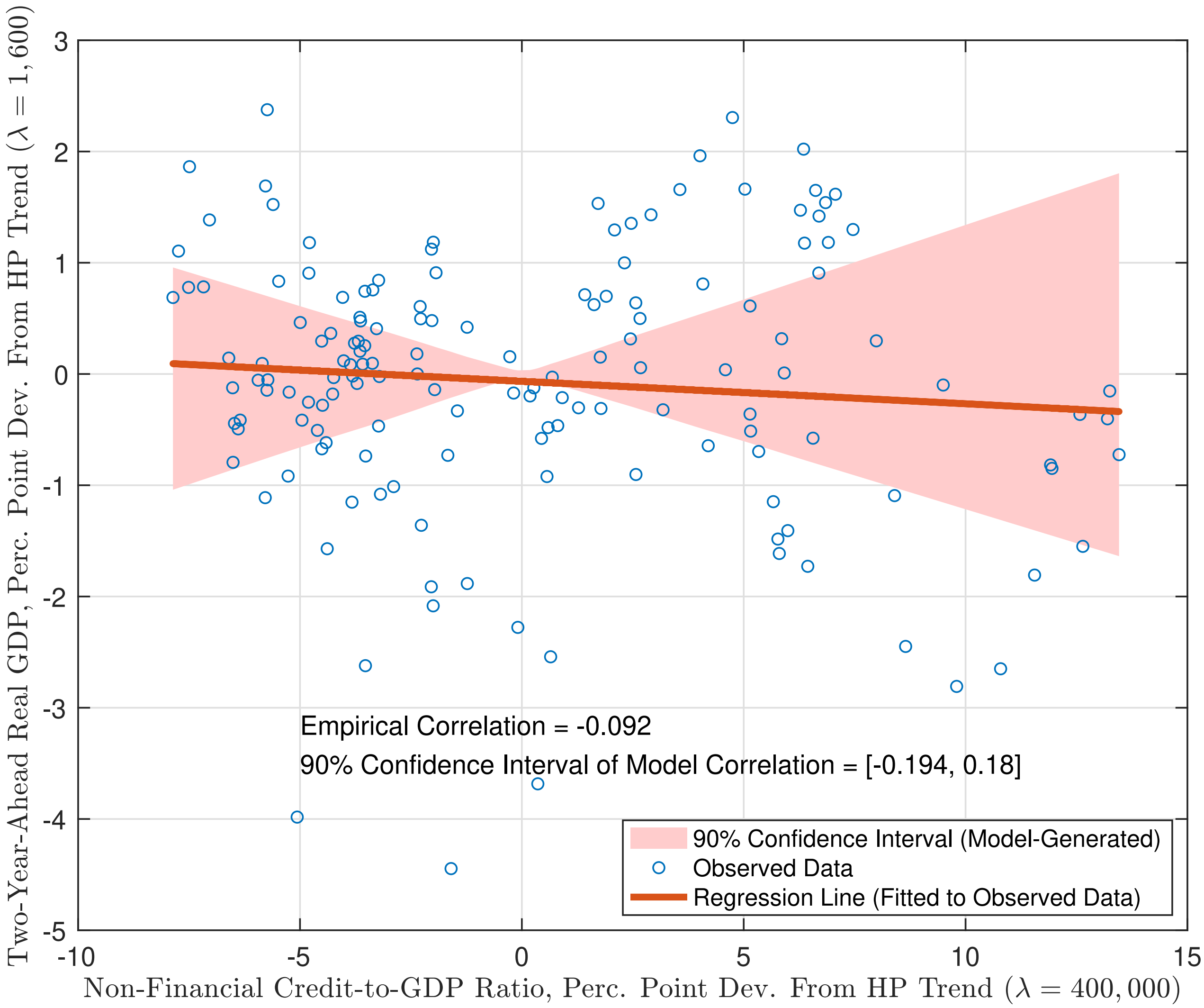
Note: This table reports the performance of various policy rules in our calibration. The first column lists the variables in the rule; the second column shows the R-squared for the regression of the Ramsey policy settings on the listed endogenous variables in the rule (and a constant); the third column shows the slope regression coefficients; the fourth and fifth columns show the average level of deposits and the average number of risk-taking quarters per 100 years when the steady state capital requirement is 10.1 percent; the sixth and seventh columns show the same statistics when the steady-state capital requirement is raised to 10.5 percent; the eighth and ninth columns show the same statistics when the steady state capital requirement is raised to 11 percent. Convergence problems occur when the model economy falls into a risk-taking trap, one in which a risk-taking crisis is so long that our solution algorithm fails.

Table 6: The Efficiency of Static Buffers

Static Buffer	Number of quarters with excessive risk taking (per 100 years)	Average deposit
10 bp	210.8	7.678
20 bp	172.0	9.216
30 bp	140.8	10.479
40 bp	108.8	11.784
50 bp	79.2	12.920
100 bp	6.8	15.805
150 bp	0	15.991
Optimal Rule	**0**	**16.241**

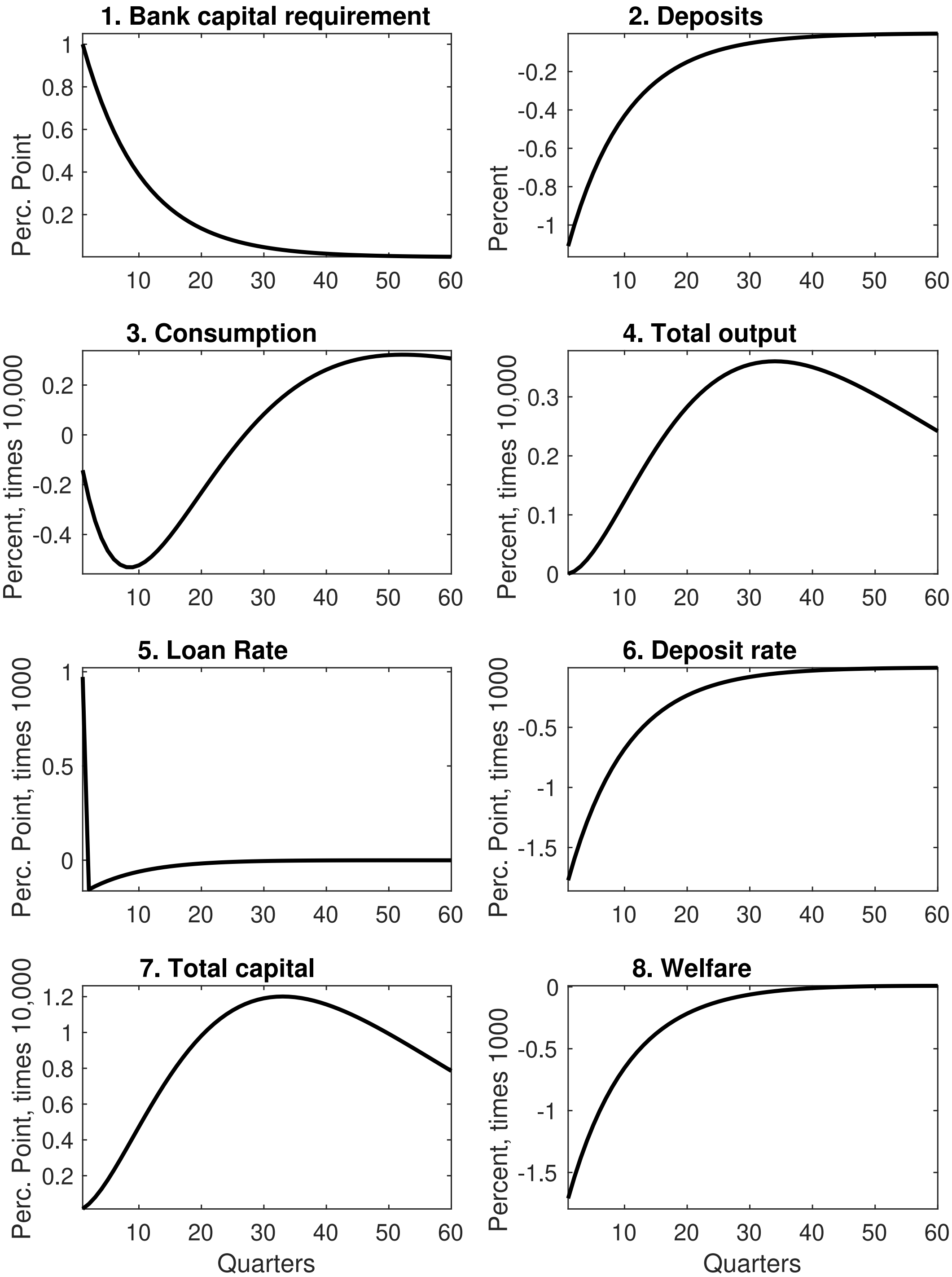
Note: This table reports the performance of static buffers in our calibration. The first column lists the proposed sizes of static buffers; the second and third columns show the average number of risk-taking quarters per 100 years and the average level of deposits when the steady state capital requirement is 10 percent plus the buffer. The last row gives the performance measures achieved by the Ramsey planner.

Figure 1: **Elevated Credit Predicts Lower GDP, Model and Data**



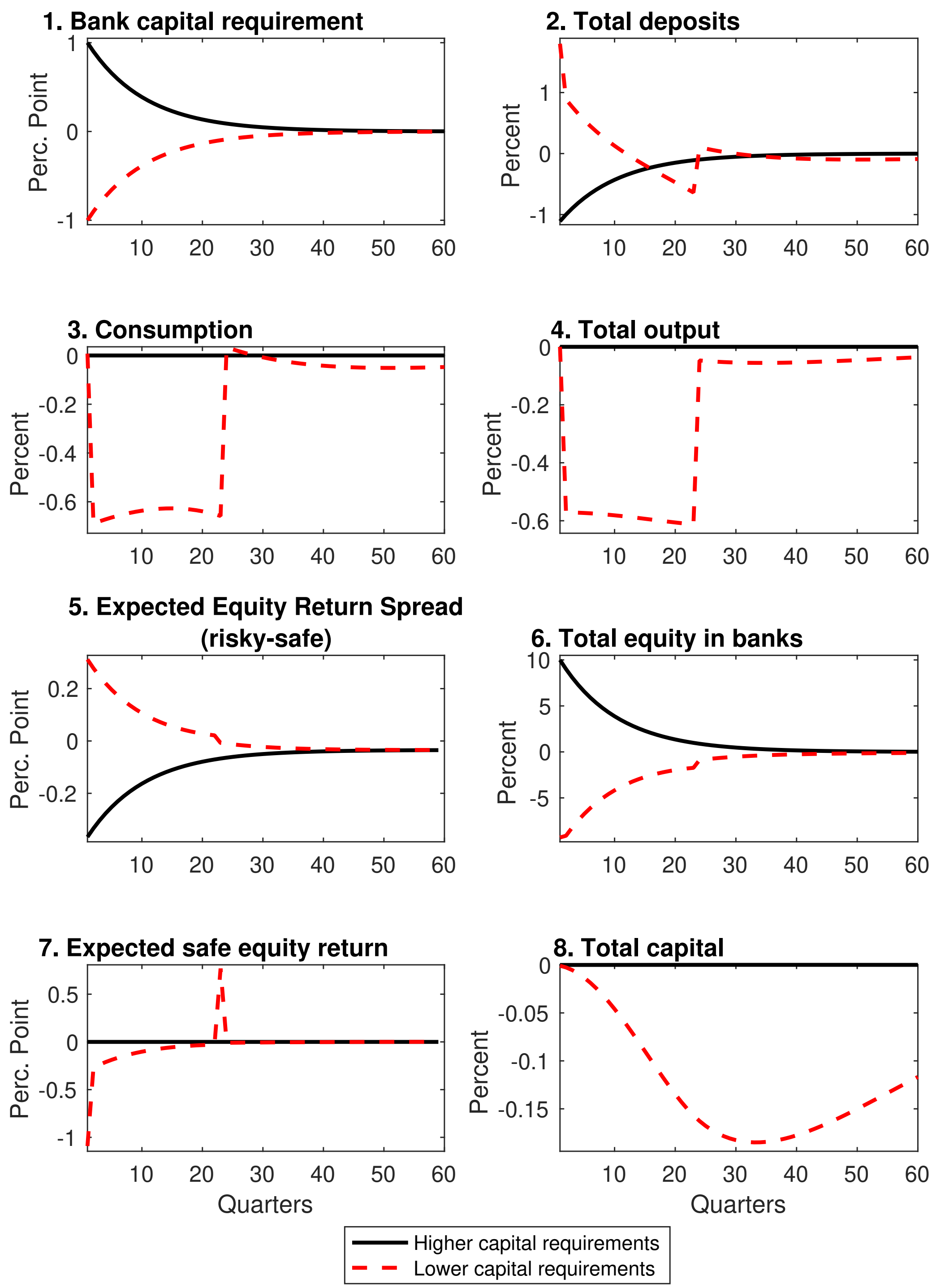
Note: This figure plots the empirical and model relationships between the ratio of non-financial credit to GDP and GDP two years ahead using data for the United States from the first quarter of 1980 through the fourth quarter of 2019. The blue open circles show the observed data. The solid line shows the regression between GDP two years ahead and the credit-to-GDP ratio from the data. The shaded area denotes a 90 percent confidence interval for the same regression slope coefficient using the model simulated data (1,000 samples of 176 observations—the same length as our observed sample). The credit-to-GDP ratio was detrended with a Hodrick-Prescott filter using a coefficient of 400,000. GDP was detrended with a Hodrick-Prescott filter using a coefficient of 1,600. The measure of real GDP (in chained (2012) dollars) comes from the National Income and Product Accounts (NIPA) of the Bureau of Economic Analysis (BEA) Table 1.1.6. For the credit-to-GDP ratio, the credit measure is for private non-financial credit, sourced from the Bank for International Settlements [CRDQUSAPABIS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CRDQUSAPABIS; for GDP we use the nominal volume from BEA NIPA Table 1.1.5.

Figure 2: **Higher Capital Requirements Reduce Deposits Leaving Little Imprint on the Rest of the Economy**



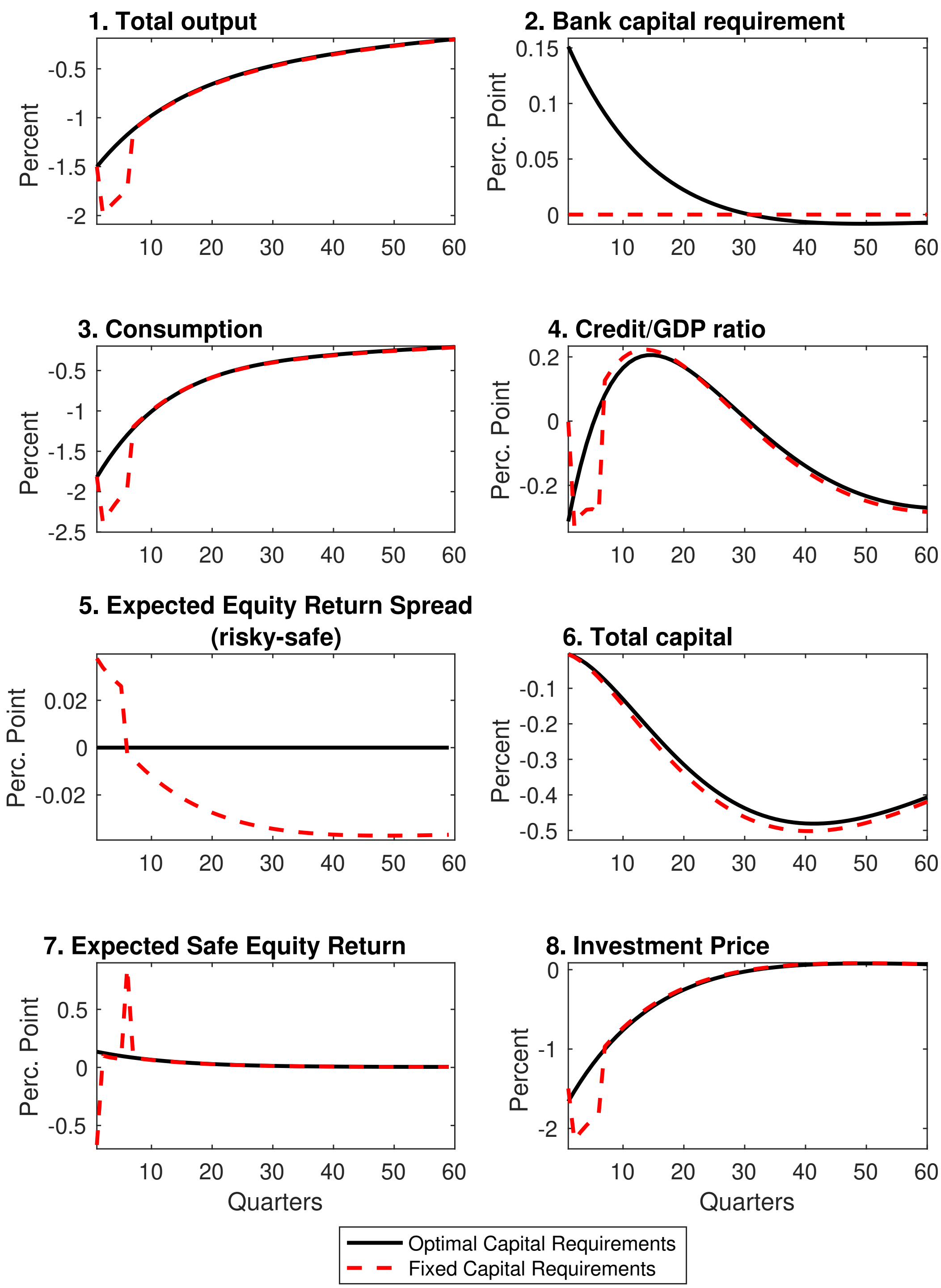
Note: This figure plots the responses of variables (in deviation from the steady state) to a one percentage point increase in the capital requirement. The shock follows an AR(1) process with a persistence parameter of 0.9.

Figure 3: **Asymmetric Effects of Higher and Lower Capital Requirements: Decreasing Capital Requirements Can Lead to a Banking Crisis.**



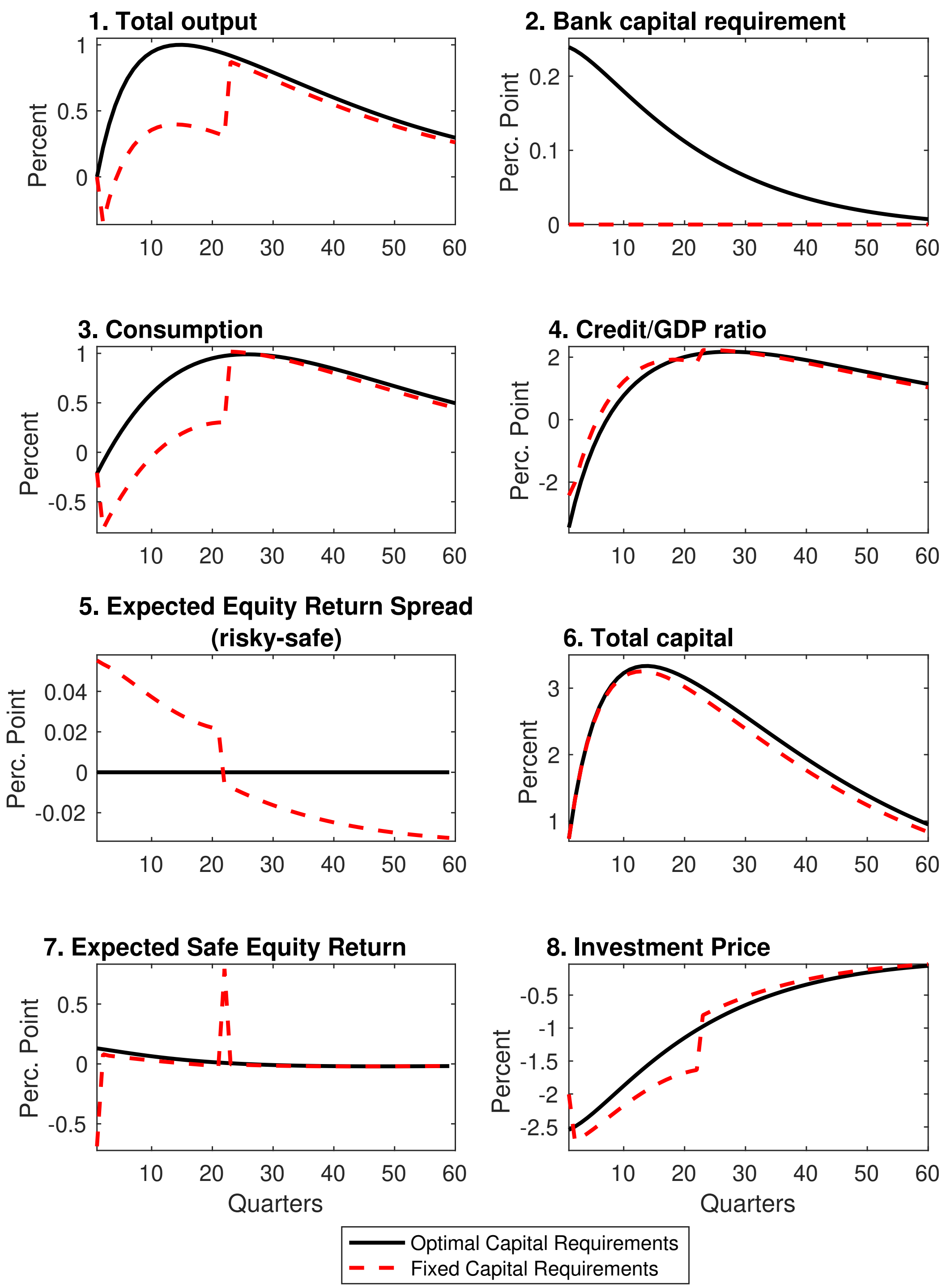
Note: This figure plots the responses of variables (in deviation from the steady state) to a one percentage point change in the capital requirement. The shock follows an AR(1) process with a persistence parameter of 0.9. The solid line shows the responses to a one percentage point rise in the capital requirement. The dashed line shows the responses to a one percentage point fall in the capital requirement.

Figure 4: **After a Negative TFP Shock, the Optimal Ramsey Policy Increases Bank Capital Requirements in a Recession**



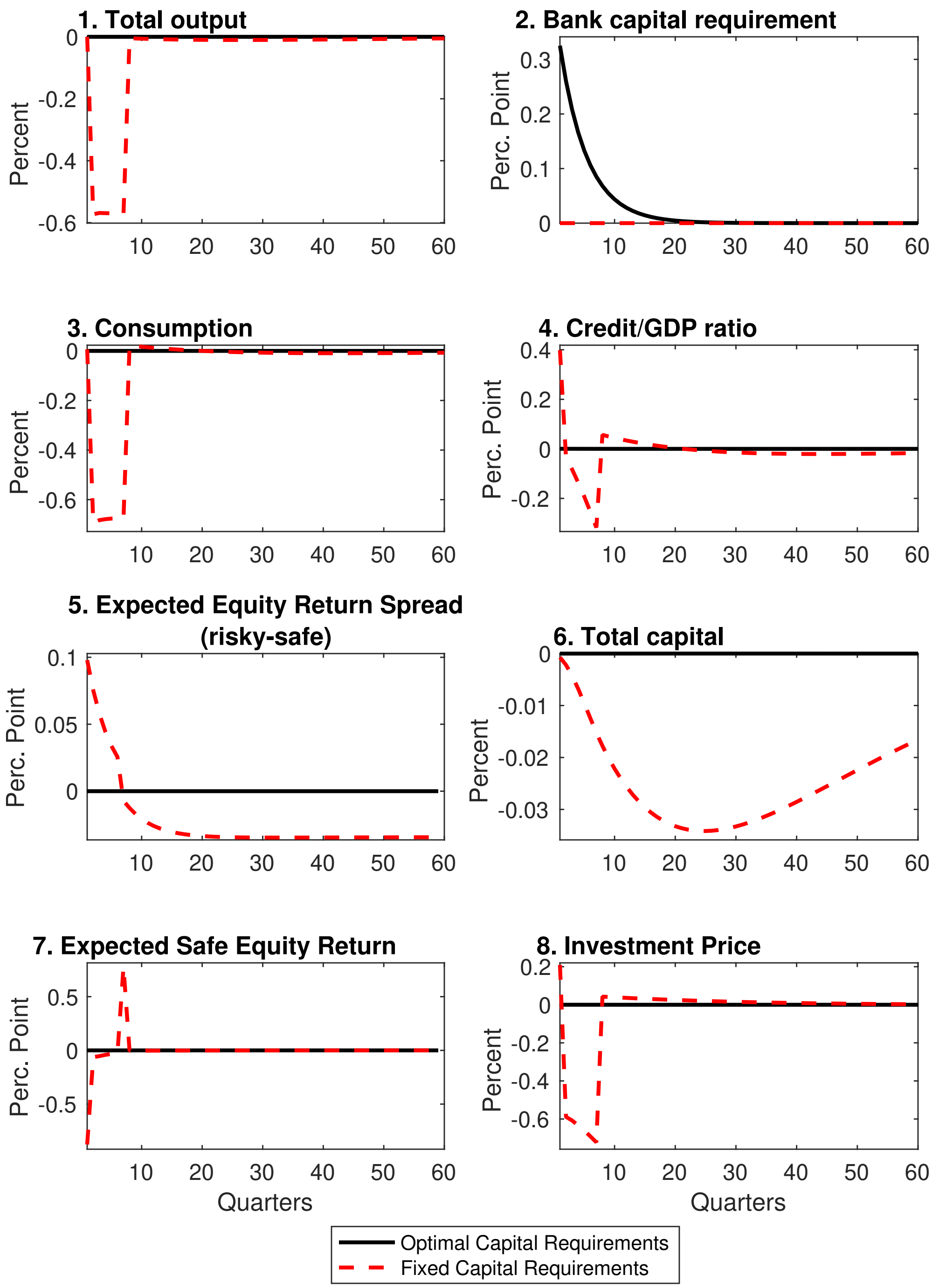
Note: This figure plots the responses of variables (in deviation from the steady state) to a 1.5 percent fall in A_t, total factor productivity. The shock follows an AR(1) process with a persistence parameter of 0.95. The solid line shows the optimal response of the capital requirement that solves the Ramsey problem. The dashed line shows what would happen if capital requirement were to be held constant at its steady state value.

Figure 5: **After a Positive Investment Shock, the Optimal Ramsey Policy Increases Bank Capital Requirements in an Expansion**



Note: This figure plots the responses of variables (in deviation from the steady state) to a positive η_t shock. The shock follows an AR(1) process with a persistence parameter of 0.8. The shock is sized to lead to an increase in output by 1% at its peak. The solid line shows the optimal response of the capital requirement that solves the Ramsey problem. The dashed line shows what would happen if capital requirement were to be held constant at its steady state value.

Figure 6: **After a Shock that Increases the Volatility of the Returns from Risky Firms, the Optimal Ramsey Policy Increases Capital Requirements Neutralizing the Macroeconomic Impact of the Shock**



Note: This figure plots the responses of variables (in deviation from the steady state) to a 15 basis point rise in the standard deviation of the idiosyncratic shock to the productivity of risky firms, τ_t. The shock follows an AR(1) process with a persistence parameter of 0.8. The solid line shows the optimal response of the capital requirement that solves the Ramsey problem. The dashed line shows what would happen if capital requirement were to be held constant at its steady state value.

10 References:

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

Angeloni, I., & Faia, E. (2013). Capital regulation and monetary policy with fragile banks. *Journal of Monetary Economics*, 60 (3), 311–324.

Barr, M. S. (2023). Holistic capital review, speech delivered at the Bipartisan Policy Center, Washington, D.C.

Basel Committee on Banking Supervision (2010). Guidance for national authorities operating the countercyclical capital buffer, Bank for International Settlements.

Begenau, J. (2020). Capital requirements, risk choice, and liquidity provision in a business-cycle model. *Journal of Financial Economics*, 136 (2), 355-378.

Begenau, J. & Landvoigt, T. (2022). Financial Regulation in a Quantitative Model of the Modern Banking System, *The Review Of Economic Studies*, 89(4), 1748–1784.

Boissay, F., Collard, F., & Smets, F. (2016). Booms and banking crises. *Journal of Political Economy,* 124(2), 489-538.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2018). Diagnostic expectations and credit cycles. *The Journal of Finance*, 73(1), 199-227.

Borio, C. & Lowe, P. (2002). Assessing the risk of banking crises, *BIS Quarterly Review*, Bank for International Settlements.

Borio, C. & Lowe, P. (2004). Securing sustainable price stability: should credit come back from the wilderness? BIS Working Papers No 157.

Brunnermeier, M. K., & Sannikov, Y. (2014). A macroeconomic model with a financial sector. *American Economic Review*, 104(2), 379-421.

Christiano, L. J., Motto, R., & Rostagno, M. (2010). Financial factors in economic fluctuations. ECB Working Paper No. 1192.

Collard, F., Dellas H, Diba B, & Loisel O. (2017). Optimal monetary and prudential policies. *American Economic Journal: Macroeconomics*, 9 (1), 40–87.

Davydiuk, T. (2017). Dynamic bank capital requirements. Mimeo, SSRN.

Faria-e-Castro, M. (2019). A quantitative analysis of countercyclical capital buffers. FRB St. Louis Working Paper, (2019-8).

Federal Register (2023). Notice of proposed rulemaking, pages 64028-64343.

Friedman, M. (1966). Interest rates and the demand for money. *The Journal of Law and Economics*, 9, 71-85.

Friedman, M., & Schwartz, A. J. (1963). *A monetary history of the United States, 1867-1960.* Princeton University Press, Princeton, NJ.

Gertler, M., & Kiyotaki, N. (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.F., Grotteria, M. & Wachter J. (2023). Foreseen Risks, *Journal of Economic Theory,* 212, 105706.

Gorton, G., & Metrick, A. (2013). Securitization. In *Handbook of the Economics of Finance* (Vol. 2, pp. 1-70). Elsevier.

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

He, Z., & Krishnamurthy, A. (2012). A model of capital and crises. *The Review of Economic Studies,* 79(2), 735-777.

Jordà, Ò., Schularick, M., & Taylor, A. M. (2017). Macrofinancial history and the new business cycle facts. *NBER Macroeconomics Annual*, 31(1), 213-263.

Malherbe, F. (2020). Optimal capital requirements over the business and financial cycles. *American Economic Journal: Macroeconomics*, 12 (3), 139-74.

Martinez-Miera, D., & Suarez, J. (2014). A macroeconomic model of endogenous systemic risk taking. mimeo.

Mendoza, E. G. (2010). Sudden stops, financial crises, and leverage. *American Economic Review*, 100(5), 1941-1966.

Mian, A., Sufi, A., & Verner, E. (2017). Household debt and business cycles worldwide. *The Quarterly Journal of Economics*, 132(4), 1755-1817.

Mian, A., Sufi, A., & Verner, E. (2020). How does credit supply expansion affect the real economy? the productive capacity and household demand channels. *Journal of Finance*, 75 (2), 949-994.

Panscott, A. & Robatto, R. The Effects of Capital Requirements on Good and Bad Risk-Taking, *The Review of Financial Studies*, 36(2), 733–774,

Schularick, M., & Taylor, A. M. (2012). Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008. *American Economic Review*, 102, 1029-1061.

Stein, J. C. (2012). Monetary policy as financial stability regulation. *The Quarterly Journal of Economics*, 127(1), 57-95.

Van den Heuvel, S. J. (2008). The welfare cost of bank capital requirements. *Journal of Monetary Economics*, 55 (2), 298–320.

Williams, John (2023). Measuring the Natural Rate of Interest: Past, Present, and Future. Speech, May 19; available on the FRBNY web page.

Online Appendix for
"A Static Capital Buffer is Hard To Beat"

A The Bank's Problem

A.1 Baseline: First-Order Conditions

Substituting $d_t = l_t - e_t$ into equation (21) and writing $\mathrm{d}G(\varepsilon_{t+1})$ explicitly turn the objective into:

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

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 $\varepsilon^*_{t+1} = \left(\frac{R^d_t + f - R^s_{t+1}}{\sigma_t} - \frac{R^d_t e_t}{\sigma_t l_t} \right) Q_t$ is the shield of limited liability. Note that we expressed ε^*_{t+1} from $\left(R^s_{t+1} + \sigma_t \frac{\varepsilon^*_{t+1}}{Q_t} \right) l_t - R^d_t \left(l_t - e_t \right) - f l_t = 0$ 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:

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

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

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

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

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

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 we omit the stochastic discount factor and the expectation operator in writing up the expressions of the next integrals. We include those terms in the final exposition.

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

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

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

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

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

Break the calculation of the integral into two parts.

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

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 $\sqrt{2}\tau$.

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

$$\begin{aligned}
&\frac{\sqrt{2}\tau}{\sqrt{\pi}}\int\limits_{\frac{\left(R_t^d(l_t-e_t)+fl_t-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}^{\infty} v\exp\left(-v^2\right)\,\mathrm{d}v - \frac{\xi}{\sqrt{\pi}}\int\limits_{\frac{\left(R_t^d(l_t-e_t)+fl_t-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}^{\infty}\exp\left(-v^2\right)\,\mathrm{d}v = \\
&-\frac{\sqrt{2}\tau}{2\sqrt{\pi}}\exp\left(-v^2\right)\Bigg|_{\frac{\left(R_t^d(l_t-e_t)+fl_t-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}^{\infty} - \frac{\xi}{\sqrt{\pi}}\left[\int\limits_0^{\infty}\exp\left(-v^2\right)\,\mathrm{d}v - \int\limits_0^{\frac{\left(R_t^d(l_t-e_t)+fl_t-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}\exp\left(-v^2\right)\,\mathrm{d}v\right] = \\
&0 + l_t\frac{\tau}{\sqrt{2\pi}}\exp\left(-\left(\frac{\left(R_t^d\left(l_t-e_t\right)+fl_t-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)^2\right) - \\
&\frac{\xi}{\sqrt{\pi}}\left[\frac{\sqrt{\pi}}{2}\mathrm{erf}(\infty) - \frac{\sqrt{\pi}}{2}\mathrm{erf}\left(\frac{\left(R_t^d\left(l_t-e_t\right)+fl_t-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right] = \\
&\frac{\tau}{\sqrt{2\pi}}\exp\left(-\left(\frac{\left(R_t^d\left(l_t-e_t\right)+fl_t-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)^2\right) - \frac{\xi}{2}\left[1-\mathrm{erf}\left(\frac{\left(R_t^d\left(l_t-e_t\right)+fl_t-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right],
\end{aligned}$$

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

Let us express $\int\limits_{\left(\frac{R_t^d+f-R_{t+1}^s}{\sigma_t}-\frac{R_t^d e_t}{\sigma_t l_t}\right)Q_t}^{\infty}\left(\frac{1}{\sqrt{2\pi\tau^2}}\exp\left(-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}\right)\right)\mathrm{d}\varepsilon_{t+1}$ in terms of the error function. Again, use the transformation $v = \frac{\varepsilon_{t+1}+\xi}{\sqrt{2}\tau}$ or $\varepsilon_{t+1} = v\sqrt{2}\tau - \xi$

$$\begin{aligned}
&\int\limits_{\frac{\left(R_t^d(l_t-e_t)+fl_t-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}^{\infty}\frac{\sqrt{2}\tau}{\sqrt{2\pi\tau^2}}\exp\left(-v^2\right)\,\mathrm{d}v = \frac{1}{\sqrt{\pi}}\int\limits_{\frac{\left(R_t^d(l_t-e_t)+fl_t-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}^{\infty}\exp\left(-v^2\right)\,\mathrm{d}v = \\
&\frac{1}{2}\left(1-\mathrm{erf}\left(\frac{\left(R_t^d\left(l_t-e_t\right)+fl_t-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right).
\end{aligned}$$

Therefore,

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

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

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

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

In sum, the FOCs can be written as follows:

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

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

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

$$\beta E_t \frac{\lambda_{ct+1}}{\lambda_{ct}} R^{e,i}_{t+1} = 1 - \frac{\zeta^i_t}{\lambda_{ct}},$$

where $i \in \{s, r\}$ denotes the type of equity. In this expression, substitute eq. (A.2) for 1. Therefore,

$$E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[R^d_t \frac{1}{2} \left(1 - \operatorname{erf}\left(\frac{\left(R^d_t \left(1 - \frac{e^i_t}{l^i_t} \right) + f - R^s_{t+1} \right) Q_t + \xi \sigma^i_t}{\sigma^i_t \sqrt{2}\tau} \right) \right) \right] - R^{e,i}_{t+1} \right\} - \frac{\zeta^i_t}{\lambda_{ct}} + \chi^i_{1t} = 0. \tag{A.3}$$

Since the range of the erf function is between -1 and 1, i.e.$-1 \leq \operatorname{erf}(x) \leq 1$, we know that the following expression is between Ψ^*_1 and Ψ^*_2:

$$\Psi^*_1 \leq E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[R^d_t \frac{1}{2} \left(1 - \operatorname{erf}\left(\frac{\left(R^d_t \left(1 - \frac{e^i_t}{l^i_t} \right) + f - R^s_{t+1} \right) Q_t + \xi \sigma^i_t}{\sigma^i_t \sqrt{2}\tau} \right) \right) - R^{e,i}_{t+1} \right] \right\} \leq \Psi^*_2,$$

where

$$\begin{aligned}\Psi^*_1 &= E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[0 - R^{e,i}_{t+1} \right] \right\},\\ \Psi^*_2 &= E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[R^d_t - R^{e,i}_{t+1} \right] \right\}.\end{aligned}$$

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

$$E_t \left\{ \beta \lambda_{ct+1} \left[R^d_t - R^{e,i}_{t+1} \right] \right\} = -\varsigma_0 D_t^{-\varsigma_d} + \zeta^i_t.$$

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

Rewriting eq. (A.3)

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

Combine it with $\Psi_2^* < \frac{\zeta_t^i}{\lambda_{ct}}$ to find

$$\frac{\zeta_t^i}{\lambda_{ct}} - \chi_{1t} < \Psi_2^* < \frac{\zeta_t^i}{\lambda_{ct}}.$$

Hence, $\chi_{1t}^i > 0$. $\square$

A.3 Combined First-Order Conditions

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

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:

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

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) $\left[\left(R^s_{t+1}+\frac{\sigma_t}{Q_t}\right)l_t - R^d_t\left(l_t-e_t\right)-fl_t\right]-\int R^e_{t+1,b}(b)\cdot e_t=0$, 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 $R^e_{t+1,b}=0$. Therefore,

$$R^e_{t+1}=\int\limits_{\left(\frac{R^d_t(1-\gamma_t)+f-R^s_{t+1}}{\sigma_t}\right)Q_t}^{\infty}\frac{\left(\left(R^s_{t+1}+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}\right)l_t-R^d_t d_t-fl_t\right)\frac{1}{\sqrt{2\pi\tau^2}}\exp\left(-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}\right)\mathrm{d}\varepsilon_{t+1}}{e_t}+$$

$$\int\limits_{-\infty}^{\left(\frac{R^d_t(1-\gamma_t)+f-R^s_{t+1}}{\sigma_t}\right)Q_t}0\cdot\frac{1}{\sqrt{2\pi\tau^2}}\exp\left(-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}\right)\mathrm{d}\varepsilon_{t+1}=$$

$$\frac{1}{e_t}\int\limits_{\left(\frac{R^d_t(1-\gamma_t)+f-R^s_{t+1}}{\sigma_t}\right)Q_t}^{\infty}\left(R^s_{t+1}l_t-R^d_t d_t-fl_t\right)\frac{1}{\sqrt{2\pi\tau^2}}\exp\left(-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}\right)\mathrm{d}\varepsilon_{t+1}+$$

$$\frac{1}{e_t}\int\limits_{\left(\frac{R^d_t(1-\gamma_t)+f-R^s_{t+1}}{\sigma_t}\right)Q_t}^{\infty}\sigma_t\frac{\varepsilon_{t+1}}{Q_t}l_t\frac{1}{\sqrt{2\pi\tau^2}}\exp\left(-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}\right)\mathrm{d}\varepsilon_{t+1}=$$

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

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

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

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

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

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

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

$$\Omega\left(\mu_t, \sigma_t;\, l_t,\, d_t,\, e_t\right) =$$
$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\int_{\left(\frac{R^d_t(l_t-e_t)+fl_t}{\sigma_t l_t}-\frac{R^s_{t+1}}{\sigma_t}\right)Q_t}^{\infty}\left(\left(R^s_{t+1}+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}\right)l_t-R^d_t\left(l_t-e_t\right)-fl_t\right)\frac{1}{\sqrt{2\pi\tau^2}}\exp\left(-\frac{\left(\varepsilon_{t+1}+\xi\right)^2}{2\tau^2}\right)\mathrm{d}\varepsilon_{t+1}\right]=$$

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

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\frac{\sigma_t l_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}\exp\left(-\left(\frac{\left(R^d_t\left(l_t-e_t\right)+fl_t-R^s_{t+1}l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)^2\right)+\right.\right.$$
$$\left.\left.\frac{\left(R^s_{t+1}l_t-R^d_t\left(l_t-e_t\right)-fl_t-\frac{\sigma_t\xi}{Q_t}l_t\right)}{2}\left[1-\mathrm{erf}\left(\frac{\left(R^d_t\left(l_t-e_t\right)+fl_t-R^s_{t+1}l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right]\right]\right\}.$$

B The Non-Financial Firm's Problem

B.1 Safe firms

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

$$\pi^s_{t+1} = y^s_{t+1} + (1-\delta)Q_t k^s_{t+1} - W_{t+1}h^s_{t+1} - R^s_{t+1}l^{f,s}_t.$$

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

$$\max_{l^{f,s}_t, k^s_{t+1}} E_t \left\{ \max_{h^s_{t+1}} \pi^s_{t+1} \right\}.$$

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

$$W_{t+1} = \frac{\partial y^s_{t+1}}{\partial h^s_{t+1}} = (1-\alpha)\frac{y^s_{t+1}}{h^s_{t+1}} = (1-\alpha) A_{t+1}\left(\frac{k^s_{t+1}}{h^s_{t+1}}\right)^{\alpha}, \tag{B.1}$$

$$h^s_{t+1} = (1-\alpha)\frac{y^s_{t+1}}{W_{t+1}} = (1-\alpha)\frac{A_{t+1}\left(k^s_{t+1}\right)^{\alpha}\left(h^s_{t+1}\right)^{1-\alpha}}{W_{t+1}}. \tag{B.2}$$

Accordingly, the safe firm's Lagrangian is:

$$\begin{aligned}\mathcal{L}^{\text{safe}} = &E_t\left\{A_{t+1}\left(k^s_{t+1}\right)^{\alpha}\left(h^s_{t+1}\right)^{1-\alpha} + (1-\delta)Q_{t+1}k^s_{t+1} - W_{t+1}h^s_{t+1} - R^s_{t+1}l^{f,s}_t\right\} + \\ &\lambda^s_{ht}E_t\left\{(1-\alpha)\frac{A_{t+1}\left(k^s_{t+1}\right)^{\alpha}\left(h^s_{t+1}\right)^{1-\alpha}}{W_{t+1}} - h^s_{t+1}\right\} + \lambda^s_{lt}\left(l^{f,s}_t - Q_t k^s_{t+1}\right).\end{aligned}$$

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

$$\frac{\partial \mathcal{L}^{\text{safe}}}{\partial l^{f,s}_t} = -E_t\left[R^s_{t+1}\right] + \lambda^s_{lt} = 0,$$

$$\frac{\partial \mathcal{L}^{\text{safe}}}{\partial k^s_{t+1}} = E_t\left[\alpha\frac{y^s_{t+1}}{k^s_{t+1}} + (1-\delta)Q_{t+1}\right] + \lambda^s_{ht}(1-\alpha)\alpha E_t\left[\frac{A_{t+1}}{W_{t+1}}\left(\frac{k^s_{t+1}}{h^s_{t+1}}\right)^{\alpha-1}\right] - \lambda^s_{lt}Q_t = 0,$$

$$\frac{\partial \mathcal{L}^{\text{safe}}}{\partial h^s_{t+1}} = (1-\alpha)\frac{A_{t+1}\left(k^s_{t+1}\right)^{\alpha}\left(h^s_{t+1}\right)^{1-\alpha}}{W_{t+1}} - W_{t+1} + \lambda^s_{ht}\left[(1-\alpha)^2\frac{A_{t+1}}{W_{t+1}}\left(\frac{k^s_{t+1}}{h^s_{t+1}}\right)^{\alpha} - 1\right] = 0.$$

Combining $\frac{\partial \mathcal{L}^{\text{safe}}}{\partial h^s_{t+1}} = 0$ with equation (B.2) yields $\lambda^s_{ht} = 0$. Then, plugging $\frac{\partial \mathcal{L}^{\text{safe}}}{\partial l^{f,s}_t} = 0$ into

$\frac{\partial \mathcal{L}^{\text{safe}}}{\partial k^s_{t+1}}$ for λ^s_{lt}, we get

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

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

$$y^s_{t+1} = \frac{\partial y^s_{t+1}}{\partial k^s_{t+1}}k^s_{t+1} + \frac{\partial y^s_{t+1}}{\partial h^s_{t+1}}h^s_{t+1} = \alpha A_{t+1}\left(\frac{k^s_{t+1}}{h^s_{t+1}}\right)^{\alpha-1}k^s_{t+1} + W_{t+1}h^s_{t+1},$$

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

$$\alpha A_{t+1}\left(\frac{k^s_{t+1}}{h^s_{t+1}}\right)^{\alpha-1}k^s_{t+1} + (1-\delta)Q_{t+1}k^s_{t+1} - R^s_{t+1}Q_t k^s_{t+1} = 0.$$

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

$$R^s_{t+1}Q_t = \alpha A_{t+1}\left(\frac{k^s_{t+1}}{h^s_{t+1}}\right)^{\alpha-1} + (1-\delta)Q_{t+1} \tag{B.3}$$

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

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

B.2 Risky Firms

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

$$\pi^r_{t+1} = y^r_{t+1} + (1-\delta)Q_t k^r_{t+1} - W_{t+1}h^r_{t+1} - R^r_{t+1}l^{f,r}_t.$$

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

$$\max_{l^{f,r}_t, k^r_{t+1}} E_t\left\{\max_{h^r_{t+1}} \pi^r_{t+1}\right\}.$$

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

$$\begin{aligned} W_{t+1} &= \frac{\partial y^r_{t+1}}{\partial h^r_{t+1}} = (1-\alpha) A_{t+1} \left(\frac{k^r_{t+1}}{h^r_{t+1}} \right)^{\alpha}, && \text{(B.4)} \\ h^r_{t+1} &= (1-\alpha) \frac{A_{t+1} \left(k^r_{t+1}\right)^{\alpha} \left(h^r_{t+1}\right)^{1-\alpha}}{W_{t+1}}. && \text{(B.5)} \end{aligned}$$

Accordingly, the risky firm's Lagrangian is:

$$\begin{aligned} \mathcal{L}^{\text{risky}} =& E_t \left[A_{t+1} \left(k^r_{t+1}\right)^{\alpha} \left(h^r_{t+1}\right)^{1-\alpha} + \varepsilon_{t+1} k^r_{t+1} + (1-\delta) Q_{t+1} k^r_{t+1} - W_{t+1} h^r_{t+1} - R^r_{t+1} l^{f,r}_t \right] + \\ & \lambda^r_{ht} E_t \left[(1-\alpha) \frac{A_{t+1} \left(k^r_{t+1}\right)^{\alpha} \left(h^r_{t+1}\right)^{1-\alpha}}{W_{t+1}} - h^r_{t+1} \right] + \lambda^r_{lt} \left(l^{f,r}_t - Q_t k^r_{t+1} \right). \end{aligned}$$

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

$$\begin{aligned} \frac{\partial \mathcal{L}^{\text{risky}}}{\partial l^{f,r}_t} &= -E_t \left[R^r_{t+1} \right] + \lambda^r_{lt} = 0, \\ \frac{\partial \mathcal{L}^{\text{risky}}}{\partial k^r_{t+1}} &= E_t \left[\alpha A_{t+1} \left(\frac{k^r_{t+1}}{h^r_{t+1}} \right)^{\alpha-1} + \varepsilon_{t+1} + (1-\delta) Q_{t+1} \right] + \\ & \quad \lambda^r_{ht} E_t \left[\alpha (1-\alpha) \frac{A_{t+1}}{W_{t+1}} \left(\frac{k^r_{t+1}}{h^r_{t+1}} \right)^{\alpha-1} \right] - \lambda^r_{lt} Q_t = 0, \\ \frac{\partial \mathcal{L}^{\text{risky}}}{\partial h^r_{t+1}} &= (1-\alpha) A_{t+1} \left(\frac{k^r_{t+1}}{h^r_{t+1}} \right)^{\alpha} - W_{t+1} + \lambda^r_{ht} \left[(1-\alpha)^2 \frac{A_{t+1}}{W_{t+1}} \left(\frac{k^r_{t+1}}{h^r_{t+1}} \right)^{\alpha} - 1 \right] = 0. \end{aligned}$$

Equation (B.4) together with $\frac{\partial \mathcal{L}^{\text{risky}}}{\partial h^r_{t+1}} = 0$ yield $\lambda^r_{ht} = 0$. Plugging $\frac{\partial \mathcal{L}^{\text{risky}}}{\partial l^{f,r}_t} = 0$ into $\frac{\partial \mathcal{L}^{\text{risky}}}{\partial k^r_{t+1}}$ for λ^r_{lt}, we get

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

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

$$\frac{k^s_{t+1}}{h^s_{t+1}} = \frac{k^r_{t+1}}{h^r_{t+1}} \qquad \text{(B.6)}$$

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 $k^i_{j,t}$ denote 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, 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.

Note 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:

$$Y_t^s = \int_{\sigma_{t-1}}^{1} \int_{0}^{1} A_t \left(k_{j,t}^i\right)^{\alpha} \left(h_{j,t}^i\right)^{1-\alpha} dj di = \int_{\sigma_{t-1}}^{1} \int_{0}^{1} F\left(k_{j,t}^i, h_{j,t}^i\right) dj di =$$

Using that the technology has Constant Returns to Scale:

$$= \int_{\sigma_{t-1}}^{1} \int_{0}^{1} \left[F_{k_{j,t}^i}\left(k_{j,t}^i, h_{j,t}^i\right) k_{j,t}^i + F_{h_{j,t}^i}\left(k_{j,t}^i, h_{j,t}^i\right) h_{j,t}^i\right] dj di =$$

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

$$= \int_{\sigma_{t-1}}^{1} \int_{0}^{1} \left[f_{k_{j,t}^i}\left(\frac{k_{j,t}^i}{h_{j,t}^i}\right) k_{j,t}^i + f_{h_{j,t}^i}\left(\frac{k_{j,t}^i}{h_{j,t}^i}\right) h_{j,t}^i\right] dj di = \text{Plugging equation (B.7)} =$$

$$= \int_{\sigma_{t-1}}^{1} \int_{0}^{1} \left[f_{k_t}\left(\frac{k_t}{h_t}\right) k_{j,t}^i + f_{h_t}\left(\frac{k_t}{h_t}\right) h_{j,t}^i\right] dj di =$$

$$f_{k_t}\left(\frac{k_t}{h_t}\right)\left[\int_{\sigma_t}^{1} \int_{0}^{1} k_{j,t}^i dj di\right] + f_{h_t}\left(\frac{k_t}{h_t}\right)\left[\int_{\sigma_t}^{1} \int_{0}^{1} h_{j,t}^i dj di\right] = f_{k_t}\left(\frac{k_t}{h_t}\right) K_t^s + f_{h_t}\left(\frac{k_t}{h_t}\right) H_t^s =$$

Since $\frac{K_{s,t}^s}{H_{s,t}^s} = \frac{K_{r,t}^s}{H_{r,t}^s} = \frac{k_t}{h_t}$,then $\frac{K_t^s}{H_t^s}\frac{h_t}{k_t} = \left(\frac{K_{s,t}^s + K_{r,t}^s}{H_{s,t}^s + H_{r,t}^s}\right)\frac{H_{r,t}^s}{K_{r,t}^s} = 1$. Therefore $\frac{K_t^s}{H_t^s} = \frac{k_t}{h_t}$.

$$= f_{K_t^s}\left(\frac{K_t^s}{H_t^s}\right) K_t^s + f_{H_t^s}\left(\frac{K_t^s}{H_t^s}\right) H_t^s = A_t \left(K_t^s\right)^{\alpha} \left(H_t^s\right)^{1-\alpha}.$$

Risky representative firm:

$$Y_t^r = \int_{0}^{\sigma_{t-1}} \int_{0}^{1} \left[A_t \left(k_{j,t}^i\right)^{\alpha} \left(h_{j,t}^i\right)^{1-\alpha} + \varepsilon_{j,t}^i k_{j,t}^i\right] dj di = \int_{0}^{\sigma_{t-1}} \int_{0}^{1} F\left(k_{j,t}^i, h_{j,t}^i\right) dj di + \int_{0}^{\sigma_{t-1}} \int_{0}^{1} \varepsilon_{j,t}^i k_{j,t}^i dj di$$

Note that the similar steps described above apply to the first term in the summation, so that $\int_0^{\sigma_{t-1}} \int_0^1 F\left(k_{j,t}^i, h_{j,t}^i\right) dj di = A_t \left(K_t^r\right)^{\alpha} \left(H_t^r\right)^{1-\alpha}$. 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.

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

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

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

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

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

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 of Van den Heuvel (2008) to encompass 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 that can drive our results.

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[37], so $\underline{\varepsilon}$ and $\bar{\varepsilon}$ can take $-\infty$ and $+\infty$, respectively. Note that G_ε need not be continuous.

Let $\hat{\varepsilon}(\sigma_t,\, R^s_{t+1}) \equiv \left(\frac{R^d_t d_t}{\sigma_t l_t} - \frac{R^s_{t+1}+f}{\sigma_t}\right) Q_t = \frac{R^d_t(1-\gamma_t)+f-R^s_{t+1}}{\sigma_t} Q_t$, where the latter equation uses the result that the capital requirement constraint always binds. It denotes the realization of the idiosyncratic shock below which the bank's net worth is negative. Let $\pi(\sigma_t,\, R^s_{t+1}) = E_\varepsilon\left[\left(\left(R^s_{t+1} - f + \frac{\sigma_t \varepsilon}{Q_t}\right) l_t - R^d_t d_t\right)^+\right]$ be a function of expected dividends (taken over the idiosyncratic shock only) under some realization of R^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

$$
\begin{aligned}
&\Pi(\sigma_t) = \int_\Omega \pi\left(\sigma_t,\, R^s_{t+1}(\omega)\right) P(d\omega) = E_t\left[\int_{\hat{\varepsilon}(\sigma_t,\, R^s_{t+1})}^{\bar{\varepsilon}} \left(\left(R^s_{t+1} - f + \frac{\sigma_t \varepsilon}{Q_t}\right) l_t - R^d_t d_t\right) dG_\varepsilon\right] = \\
&E_t\left[\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left(\left(R^s_{t+1} - f + \frac{\sigma_t \varepsilon}{Q_t}\right) l_t - R^d_t d_t\right) dG_\varepsilon\right] - E_t\left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t,\, R^s_{t+1})} \left(\left(R^s_{t+1} - f + \frac{\sigma_t \varepsilon}{Q_t}\right) l_t - R^d_t d_t\right) dG_\varepsilon\right] = \\
&E_t R^s_{t+1} l_t - R^d_t d_t - f l_t - \frac{\sigma_t \xi}{Q_t} l_t - \frac{\sigma_t l_t}{Q_t} E_t\left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t,\, R^s_{t+1})} \left(\varepsilon - \hat{\varepsilon}(\sigma_t,\, R^s_{t+1})\right) dG_\varepsilon\right] = \\
&\qquad E_t R^s_{t+1} l_t - R^d_t d_t - f l_t + \frac{l_t}{Q_t}\left(\sigma_t E_t\left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t,\, R^s_{t+1})} \left(\hat{\varepsilon}(\sigma_t,\, R^s_{t+1}) - \varepsilon\right) dG_\varepsilon\right] - \sigma_t \xi\right).
\end{aligned}
$$

Note that in the derivations above we express $\left(R^s_{t+1} - f + \frac{\sigma_t \varepsilon}{Q_t}\right) l_t - R^d_t d_t$ 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

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

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 $\hat{\varepsilon}(\sigma_t,\, R^s_{t+1}) = \frac{R^d_t(1-\gamma_t)+f-R^s_{t+1}}{\sigma_t} > 0$, so $R^s_{t+1} < R^d_t\,(1-\gamma_t)+f$,

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

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

   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 $\hat{\varepsilon}_i = \hat{\varepsilon}(\sigma_{it},\, R^s_{t+1}) \equiv \frac{R^d_t(1-\gamma_t)+f-R^s_{t+1}}{\sigma_{it}} Q_t$, for $i = 1,\ 2,\ \lambda$.

(a) $R^s_{t+1} < R^d_t(1-\gamma_t) + f$: 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})} (\hat{\varepsilon}(\sigma_{\lambda t}) - \varepsilon)\, dG_\varepsilon \right\} = \\
\lambda\sigma_{1t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon - \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \right\} + \\
(1-\lambda)\sigma_{2t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon + \int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \right\} = \\
\lambda\sigma_{1t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_1) + \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\varepsilon - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \right\} + \\
(1-\lambda)\sigma_{2t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_2) + \int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \right\} \leq \\
\lambda\sigma_{1t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_1) + \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \right\} + \\
(1-\lambda)\sigma_{2t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_2) + \int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, dG_\varepsilon \right\},
\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:

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

where we use that $\sigma_{1t} = l_t \left(R^d_t (1-\gamma_t) + f - R^s_{t+1} \right) = \sigma_{2t}\hat{\varepsilon}_2 = \sigma_{\lambda t}\hat{\varepsilon}_\lambda$ in the last equality. So,

$$
\begin{aligned}
\lambda\sigma_{1t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1) + (1-\lambda)\sigma_{2t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2) = \\
\hat{\varepsilon}_\lambda (\lambda\sigma_{1t} + (1-\lambda)\sigma_{2t}) - \left(R^d_t (1-\gamma_t) + f - R^s_{t+1} \right) (\lambda + (1-\lambda)) = \\
\sigma_{\lambda t}\hat{\varepsilon}_\lambda - \left(R^d_t (1-\gamma_t) + f - R^s_{t+1} \right) = \left(R^d_t (1-\gamma_t) + f - R^s_{t+1} \right) - \left(R^d_t (1-\gamma_t) + f - R^s_{t+1} \right) = 0.
\end{aligned}
$$

Therefore, $h(\sigma_t)$ is convex in σ_t for $R^s_{t+1} < R^d_t(1-\gamma_t) + f$.

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

$$\begin{aligned} 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\}, \end{aligned}$$

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:

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

where the last equality follows from the same reasoning employed in the previous case. Therefore, $h(\sigma_t)$ is convex in σ_t for $R^s_{t+1} > R^d_t(1-\gamma_t)+f$.

(c) $R^s_{t+1} = R^d_t(1-\gamma_t)+f$. 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 $R^s_{t+1}(\omega) \in \mathbb{R}$. Then the following function[38]:

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

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. $\square$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

F Discussion of the Excessive Risk-Taking Mechanism

Following our 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:

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

where

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

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

The first component, ω_1, represents 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, $R^s_{t+1} - R^d_t\left(1-\gamma_t\right) - f - \frac{\xi\sigma_t}{Q_t}$, 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 component, ω_2, represents the extra effect of σ_t on expected dividends owing to 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 A.1 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.

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. $R^s_{t+1} - R^d_t(1-\gamma_t) - f - \frac{\xi\sigma_t}{Q_t}$ 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, $R_{t+1}^s - R_t^d(1-\gamma_t) - f$, 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, $R_{t+1}^s < R_t^d(1-\gamma_t) + f$ 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, $-\xi < R_t^d(1-\gamma_t) + f - R_{t+1}^s < 0$ 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, $R_t^d(1-\gamma_t) + f - R_{t+1}^s < -\xi$ 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

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

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 non-bank sector which we do not model here. The risky bank that finances the maximum share of risky projects earns

$$\Omega_{t+1}^{risky} = R_{t+1}^r Q_t K_{t+1}^r.$$

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

$$\Omega_{t+1}^{safe} = R_{t+1}^{s} Q_t \left(K_{t+1} + B_t\right) - Q_t B_t R_t^B = \left(R_{t+1}^{s}\left(1 + \frac{B_t}{K_{t+1}}\right) - \frac{B_t}{K_{t+1}} R_t^B\right) Q_t K_{t+1},$$

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

$$Var_t\left(R_{t+1}^{r}\right) = Var_t\left(R_{t+1}^{s}\left(1 + \frac{B_t}{K_{t+1}}\right) - \frac{B_t}{K_{t+1}} R_t^B\right)$$

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

$$Var_t\left(R_{t+1}^{r}\right) = Var_t\left(R_{t+1}^{s}\right) + \left(\frac{\sigma_{\text{RF}}}{Q_t}\right)^2 \tau^2,$$

$$Var_t\left(R_{t+1}^{s}\left(1 + \frac{B_t}{K_{t+1}}\right) - \frac{B_t}{K_{t+1}} R_t^B\right) = \left(1 + \frac{B_t}{K_{t+1}}\right)^2 Var_t\left(R_{t+1}^{s}\right),$$

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

$$Var_t\left(R_{t+1}^{s}\right) = \alpha^2 \left(\frac{K_{t+1}}{H_{t+1}}\right)^{2\alpha-2} Var_t\left(A_{t+1}\right) + (1-\delta)^2 Var_t\left(Q_{t+1}\right) + 2\alpha \left(\frac{K_{t+1}}{H_{t+1}}\right)^{\alpha-1} (1-\delta) Cov_t\left(A_{t+1}, Q_{t+1}\right).$$

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 $\sigma_{\text{RF}} = 45\%$ from the Shared National Credit Report issued by the Fed, OCC, and FDIC.

H Sensitivity Analysis

Most of the parameters for our model are standard in the literature but there are a handful of parameters specific to the key financial friction in our model whose role in our baseline results warrants further discussion.

Our first set of sensitivity results pertains to the steady-state capital requirements. Our analysis shows that a wide range of steady-state capital requirements can be supported by setting τ, the standard deviation of the risky firm's idiosyncratic shock, or ξ, the average penalty from financing risky projects, without changing any other parameter. For our calibration, we map the choice of τ into the level of risk that a bank would face when financing a firm with pre-existing debt. We treat ξ as a free parameter to set an empirically plausible steady-state capital requirement of 10 percent. We prefer this approach to taking a strong stance on the average penalty from pursuing risky projects and using the model to support a firm estimate of an optimal capital requirement in the steady state.

In our next set of sensitivity exercises, we also explore how the same parameters that strongly influence the steady-state capital requirements, τ and ξ, can affect the size of the optimal changes in capital requirements in response to shocks. As an example, we focus on the response to total factor productivty (TFP) shocks. Intuitively, the greater size of the idiosyncratic returns to risky projects or the smaller the average penalty for risky projects, the greater is the increase in capital requirements necessary to avoid excessive risk taking in response to the same-size TFP shock.

Moving beyond the parameters τ and ξ, we discuss sensitivity to the parameterization of the curvature of deposits in the utility function. In line with related papers that have explored optimal capital requirements, we choose this curvature to imply an interest elasticity of deposit supply close to 1. We find little to no difference in our results when exploring lower values of this elasticity up to one-tenth—those are the values of the interest elasticity for deposits considered in the papers most closely related to ours. We note that in the broader literature that examines the role of monetary aggregates in the conduct of monetary policy, the value of the relevant interest rate elasticity is far from settled. Calibrating the curvature of deposits in the utility function to imply a greater interest sensitivity for the households' supply of deposits results in a greater increase in funding costs for the same-size contraction in TFP. In that case, the incentive for banks to switch to risky projects under a constant capital requirement is magnified and results in a longer permanence in the regime with excessive risk taking. Accordingly a higher interest rate elasticity for deposits also results in a greater increase in capital requirements in response to the same size change in TFP.

Our final set of sensitivity results considers an alternative calibration of the shock processes in conjunction with our exploration of the relative merits of simple rules and static capital buffers. We consider an alternative calibration that only uses the two macroeconomic shocks—TFP and ISP (investment specific). We show that in this special case, there exist simple and implementable rules that track the Ramsey policy fairly well. In this case, relative prices can be used to ameliorate the problem of sorting out the relative size of the shocks. Accordingly, in this special setting, some simple rules for setting capital requirements for banks can perform almost as well as the optimal Ramsey rule. Still, it remains the case that a small capital buffer can also perform nearly as well as the Ramsey policy.

H.1 Steady-state Capital Requirements

Figure A.1 illustrates how steady-state capital requirements depend on the choices of two parameters: 1) τ, the standard deviation of the risky firm's idiosyncratic shock, and 2) ξ, the average penalty from financing risky projects. We find that these two parameters are mainly responsible for driving the variation of steady-state capital requirements. The left subplot of Figure A.1 shows the dependence of steady-state capital requirements on τ keeping all other parameters fixed. The right subplot of Figure A.1 shows the dependence of steady-state capital requirements on ξ keeping all other parameters fixed. The encircled points in red depict our baseline calibrated values.

When τ increases, the shield of limited liability becomes more attractive as the upside potential of risky assets goes up. To prevent excessive risk taking, capital requirements rise. Therefore, the line slopes upward in the left subplot of Figure A.1. Notice that steady-state capital requirements are relatively sensitive to τ, so we can achieve a wide range of steady-state capital requirements by changing τ without needing to adjust any other parameters. When ξ increases, risky projects become less attractive. Capital requirements fall to make it possible for the economy to benefit from liquidity services without affecting risk-taking profile. Therefore, the line slopes downward in the right subplot of Figure A.1. Notice that steady-state capital requirements vary from around 5% to almost 15% when ξ lies within a relatively narrow range of one percentage point. This graphical analysis demonstrates our claim that alternative choices of τ and ξ could support a wide range of capital requirements in the steady state.

H.2 Optimal Changes in Capital Requirements

For our next set of sensitivity results we show that the same parameters that strongly influence the steady state capital requirements also affect the size of the optimal adjustments in capital requirements in response to shocks. As an example of the optimal dynamic adjustments for capital requirements, we focus on the response to TFP shocks. Figure A.2 plots the maximum adjustment (in absolute value) in the optimal capital requirements and output to the same 1.5 percent contraction in TFP as in Figure 4, described in Section 6.3 of the main text. The two subplots on the left side of Figure A.2 show the maximum responses of optimal capital requirements and output for different values of τ keeping all other parameters fixed. The other subplots on the right side of Figure A.2 show the maximum responses of optimal capital requirements and output for different values of ξ keeping all other parameters fixed. 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 4. The maximum adjustment in the optimal capital requirements is also sensitive to changes in ξ for relatively small values of ξ but then it becomes almost insensitive for higher values. At the same time, the response of output, at its trough, is not affected by different alternatives of ξ.

H.3 Curvature of Deposits in the Utility Function

Sizing the curvature of deposits in the utility function, governed by the parameter ς_d, is closely related to sizing the interest elasticity of money demand, a topic of extensive interest.[39] The debate on the relevant interest elasticity of money demand (our household supply of deposits to banks) is still far from settled. As noted in Friedman (1966), a major strand of Keynesian analysis traces the implications of assuming an elasticity of money demand with respect to the interest rate as being very high, approaching infinity (in Keynes terms, liquidity preference is, if not absolute, approximately so). By contrast, Friedman and Schwartz (1963), championed a much lower estimate of 0.15.[40] For our model in which the curvature parameter and the elasticity are the inverse of each other, these stances would map into a curvature parameter, ς_d, close to 0, on the Keynesian side and close to 7 on the Monetarist side. The more recent literature continues to showcase a wide range of stances.[41] We choose a value of ς_d of 1.1 to approximate an elasticity of 1 as in Nagel (2016)—our standard utility function has a discontinuity at 1. Empirical estimates focused

[39] Bank notes and demand deposits at banks are close substitutes and deposits are an important component of money stock measures, such as M1 and M2 in the H.6 Release of the Federal Reserve Board.

[40] See Chapter 12 of Friedman and Schwartz (1963)

[41] Reviewing some prominent recent examples, Stein (2012) uses a value for the interest elasticity of money demand approaching infinity and Christiano, Motto and Rostagno (2010) choose an elasticity very close to 0.15 (they set the parameter for the curvature of money in the utility to 7).

on the interest sensitivity of deposits, as in Begenau (2020), point to values of this elasticity very close to our choice, about 0.7 (or $\varsigma_d = 1.4$). The extensive sensitivity analysis shows that elasticity values in the range 0.15–1 would imply negligible differences for our results.

Figure A.3 considers the same 1.5 percent contraction in total factor productivity as in Figure 4, described in Section 6.3 of the main text. In Figure A.3, the two lines in each panel show responses for our baseline calibration (the solid line) and for an alternative calibration with a curvature parameter for deposits in the utility function set to $\varsigma_d = 0.001$ (the dashed line) as opposed to 1.1 under our baseline. We chose $\varsigma_d = 0.001$, corresponding to an interest rate elasticity for deposits of $\frac{1}{\varsigma_d} = 1000$, as a stand-in for an infinite interest rate elasticity.

As established in Section 6.3, a contractionary TFP shock reduces the expected returns from safe projects. With a fixed capital requirement, this contraction can push banks to engage in excessive risk taking. On the household side, the shock compresses income, but consumption does not have to fall proportionately with income, as households can reduce their supply of deposits. All else equal, this reduction in deposits pushes up the funding costs and makes the shield of limited liability even more attractive for banks, after all with that shield, banks do not have to repay depositors.

The willingness of households to vary their supply of deposits as consumption or deposit rates move is governed by the parameter ς_d. The lower this parameter, the more willing households are to adjust deposits to cushion fluctuations in consumption. Notice that from the first-order condition for the household utility-maximization problem with respect to deposits (refer to equation (E.2)), one can see that the parameter ς_d governs both the inverse elasticity with respect to the deposit rate and the sensitivity of the reaction of deposits to changes in the marginal utility of consumption, through the term λ_t. In our calibration, we pin down the parameter with empirical evidence from studies that have estimated the interest sensitivity of deposits or, more broadly, money demand.

As lower values for ς_d result in a greater increase in funding costs for the same-size contraction in technology, the incentive for banks to switch to risky projects under a constant capital requirement is magnified and results in a longer permanence in the regime with excessive risk taking. In turn, when we choose capital requirements optimally, lower values for the parameter ς_d will imply that capital requirements have to rise by more, as shown in Figure A.4.

We found no visible difference for values of ς_d even lower than 0.001—intuitively, an elasticity of 1000 is already very high. We also found that the responses to technology and other shocks are indistinguishable for our baseline calibration of $\varsigma_d = 1.1$, a numerical approximation of the *log* case, relative to 1.4, the value chosen by Begenau (2020) or relative to even 7, the value estimated by Christiano, Motto and Rostagno (2010). For these higher values of the parameter, households are already so keen to maintain a stable level of deposits that increases in the inelasticity do not produce meaningful quantitative effects. To illustrate these results, we consider additional sensitivity to a broader set of parameters than those shown in Figures A.3 and A.4. Figure A.5 reports the duration of the regime with excessive risk taking in reaction to the same 1.5 percent reduction in TFP under a range of values for ς_d, keeping capital requirements fixed. We can see that the number of excessive risk-taking episodes is not affected by ς_d when $\varsigma_d > 1$.

H.4 Alternative Shock Calibration

Here we will consider an alternative calibration that just uses the two macroeconomic shocks—TFP and ISP (investment specific). We will show that in this special case, there exist simple and implementable rules that track the Ramsey policy fairly well in this calibration. So changing the shock structure that drives the economy can radically alter the ability of simple rules to perform well. But the Basel rule does not perform well for either calibration. The same as in the main text, 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.

H.4.1 Matching Moments, Shock Processes and Variance Decompositions

We follow exactly the same SMM procedure described in main text to calibrate our model. The only difference is the number of shocks that we include in the calibration. Table A.4 describes our results of the moment-matching exercise. It shows that model moments are close to data moments. Notice that there is no discernible difference in the targeted moments compared to our benchmark calibration in Table 4. Both calibrations are very good. 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 A.2 and A.3 show the shock processes and the variance decompositions associated with the calibration considered here. 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 setting, γ.

H.4.2 Implementable Capital Buffer Rules

Table A.5 reports our results for various policy rules under this alternative calibration. The first column lists the variables in the rule; the second column gives the R-squared for the rule's regression (and a constant); the third column shows the regression coefficients; the fourth and fifth 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 sixth and seventh 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 are the same as in the main text.

Table A.5 shows that the best implementable rule has capital requirements responding to the investment price. The R-squared 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 4 and 5 show that for both of the shocks that drive the economy, the investment price falls while the Ramsey capital requirement rises. Moreover, in Table A.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. In Table A.5, the R-squared 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.

The reason for the poor performance of the Basel rule can also be seen from Figures 4 and 5 that 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 A.3, the ISP shock drives the Ramsey capital requirements, while it only explains 41 percent of the variation in the credit-to-GDP ratio. 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. All these results on the performance of the Basel rule are in line with the results described in the main text.

Table A.5 also reports the performance of a rule that focuses on GDP. That rule fares no better than the Basel rule. The R-squared is virtually zero; so it is not tracking the Ramsey policy. And the performance measures are also bad.

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

H.4.3 The Efficiency of Static Capital Buffers

Table A.6 shows the results if 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 this alternative calibration, or 11.5 percent for our baseline calibration, the results are almost as good as those achieved by the Ramsey planner. If the static capital requirement is raised to 11.5 percent, the performance measures for both calibrations are very close to the optimal ones.

Tables and Figures in the Appendix

Table A.1: Illustrating the Effects of Higher Risk on Dividends.

States of nature where	Effects on ω_1		Effects on ω_2
	$R^s_{t+1} - R^d_t\,(1-\gamma_t) - f - \frac{\xi\sigma_t}{Q_t}$	$1 - G(\varepsilon^*_{t+1})$	
$R^s_{t+1} < R^d_t\,(1-\gamma_t) + f \quad \Leftrightarrow \quad \varepsilon^*_{t+1} > 0$	$\Downarrow$	$\Uparrow$	$\Uparrow$
$R^s_{t+1} > R^d_t\,(1-\gamma_t) + f \quad \Leftrightarrow \quad \varepsilon^*_{t+1} < 0$	$\Downarrow$	$\Downarrow$	if $\varepsilon^*_{t+1} > -\xi$, then $\Uparrow\Downarrow$
			if $\varepsilon^*_{t+1} \leqslant -\xi$, then $\Uparrow$

Table A.2: **Alternative Calibration With TFP and ISP (Investment Specific) Shocks, Shock Processes**

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

Table A.3: **Alternative Calibration With TFP and ISP (Investment Specific) Shocks, 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 A.4: **Alternative Calibration With TFP and ISP (Investment Specific) Shocks, 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 A.5: **Alternative Calibration With TFP and ISP (Investment Specific) Shocks, Simple Rules**

			Static buffer = 10 basis points		Static buffer = 30 basis points	
Simple rule	R squared	Regression slope coefficient	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
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 A.6: **The Efficiency of Static Buffers Across Different Calibrations**

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

Figure A.1: **Effects of the Standard Deviation of Idiosyncratic Returns for Risky Projects (τ) and the Average Penalty on Returns for Risky Projects (ξ) on Steady-state Capital Requirements**

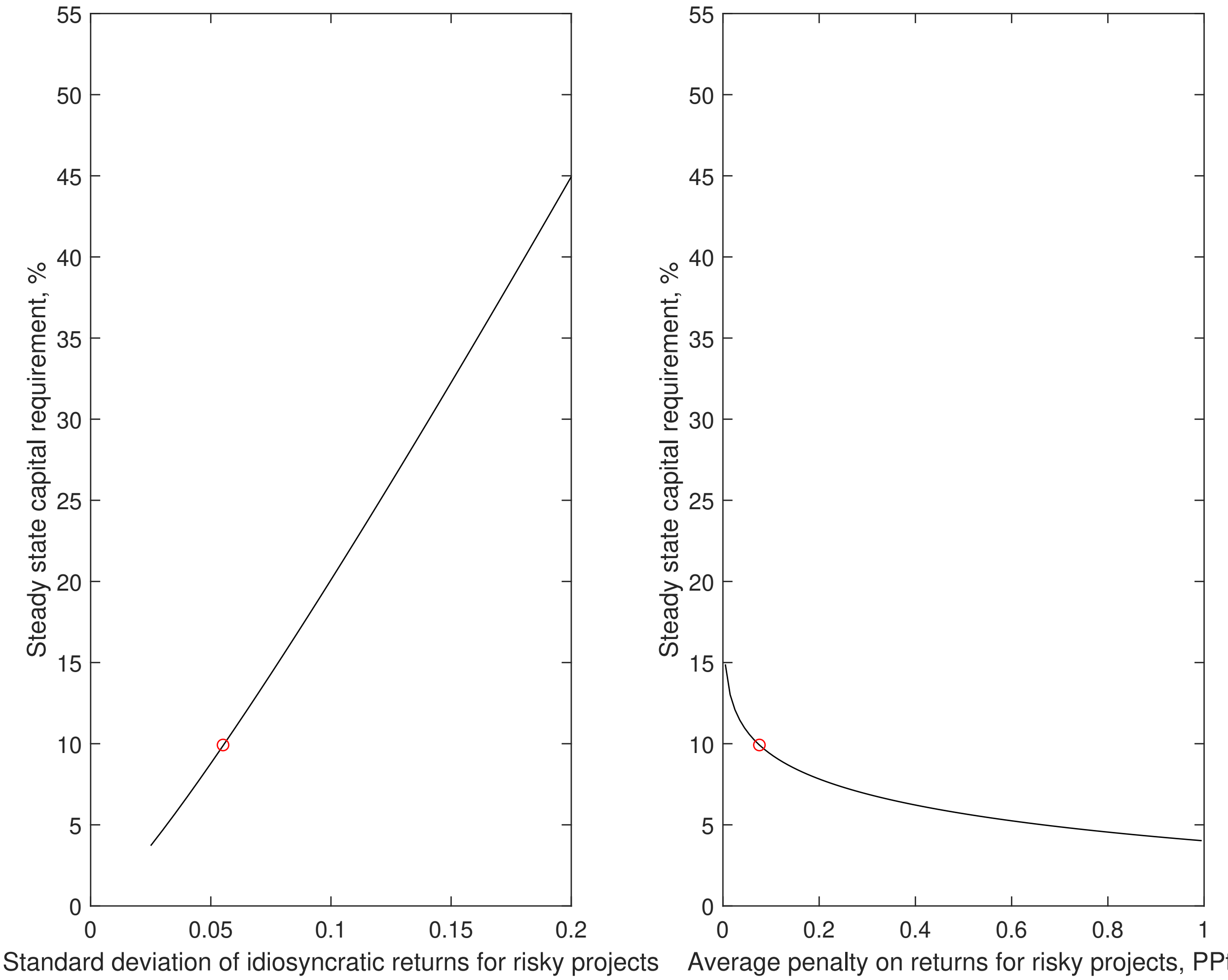


Figure A.2: **The Maximum Adjustment of Optimal Capital Requirements and Output to a Negative TFP Shock for Alternative Choices of the Parameters Governing the Standard Deviation of Idiosyncratic Returns for Risky Projects and the Average Penalty on Returns for Risky Projects (τ and ξ)**

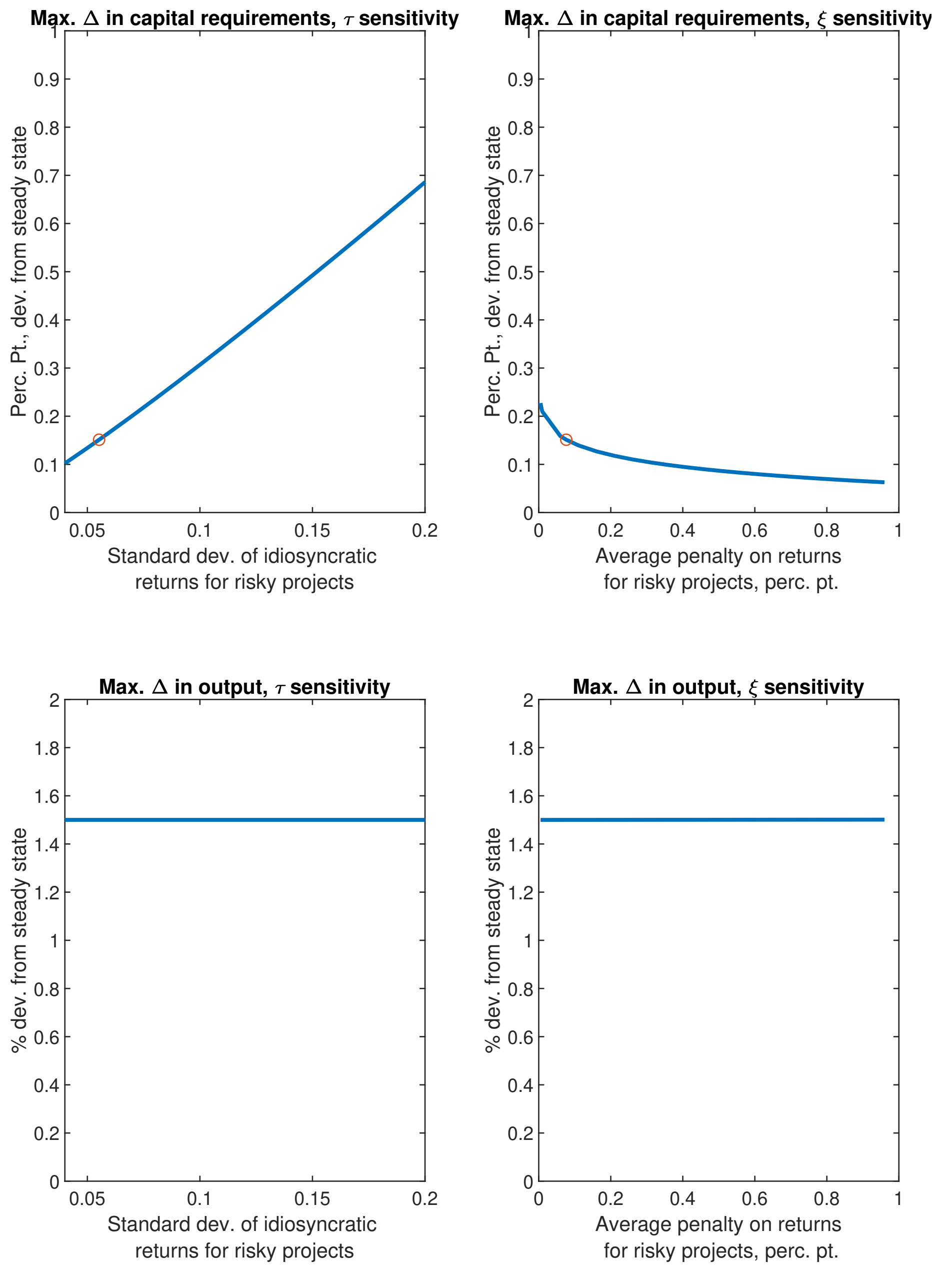


Figure A.3: **A Negative TFP Shock Under Fixed Capital Requirements for Alternative Choices of the Parameter Governing the Interest Elasticity for the Households' Supply of Deposits (elasticity $= \frac{1}{\varsigma_d}$)**

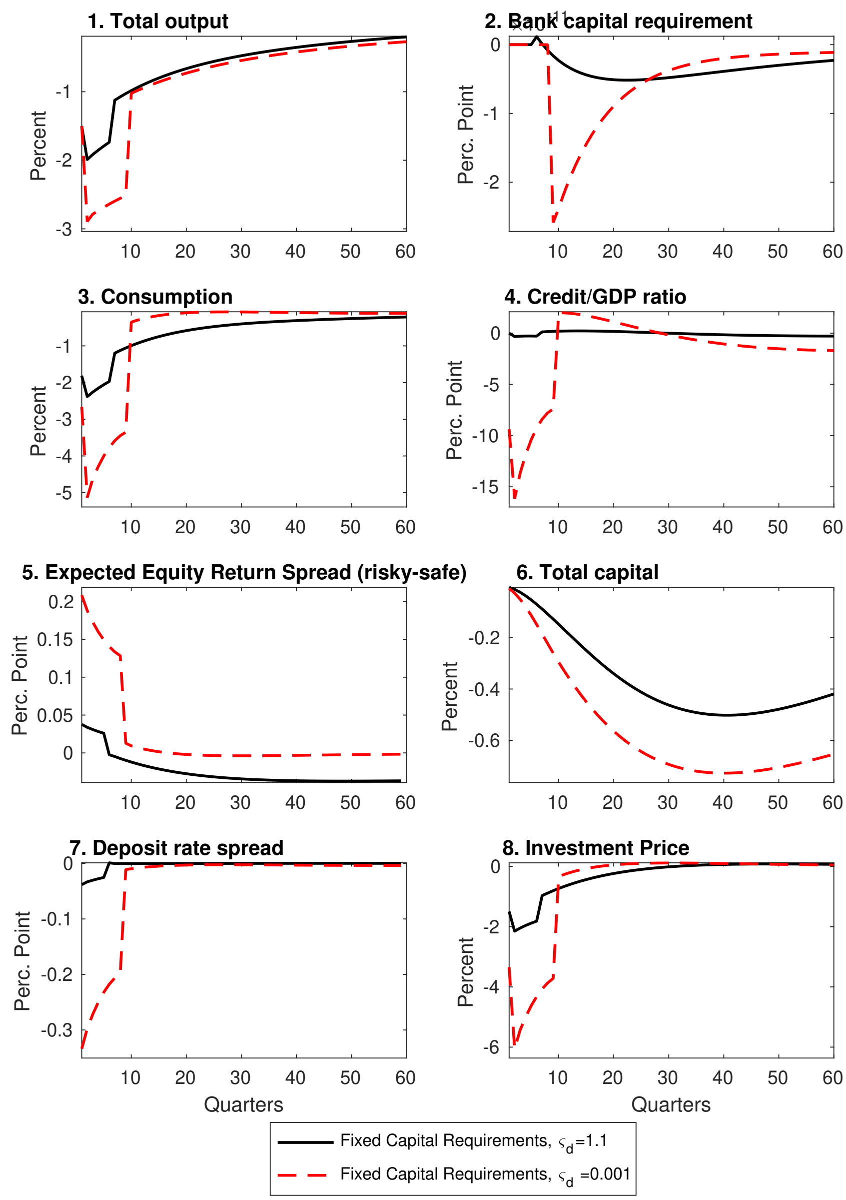


Figure A.4: **A Negative TFP Shock Under Optimal Capital Requirements for Alternative Choices of the Parameter Governing the Interest Elasticity for the Households' Supply of Deposits (elasticity** $= \frac{1}{\varsigma_d}$**)**

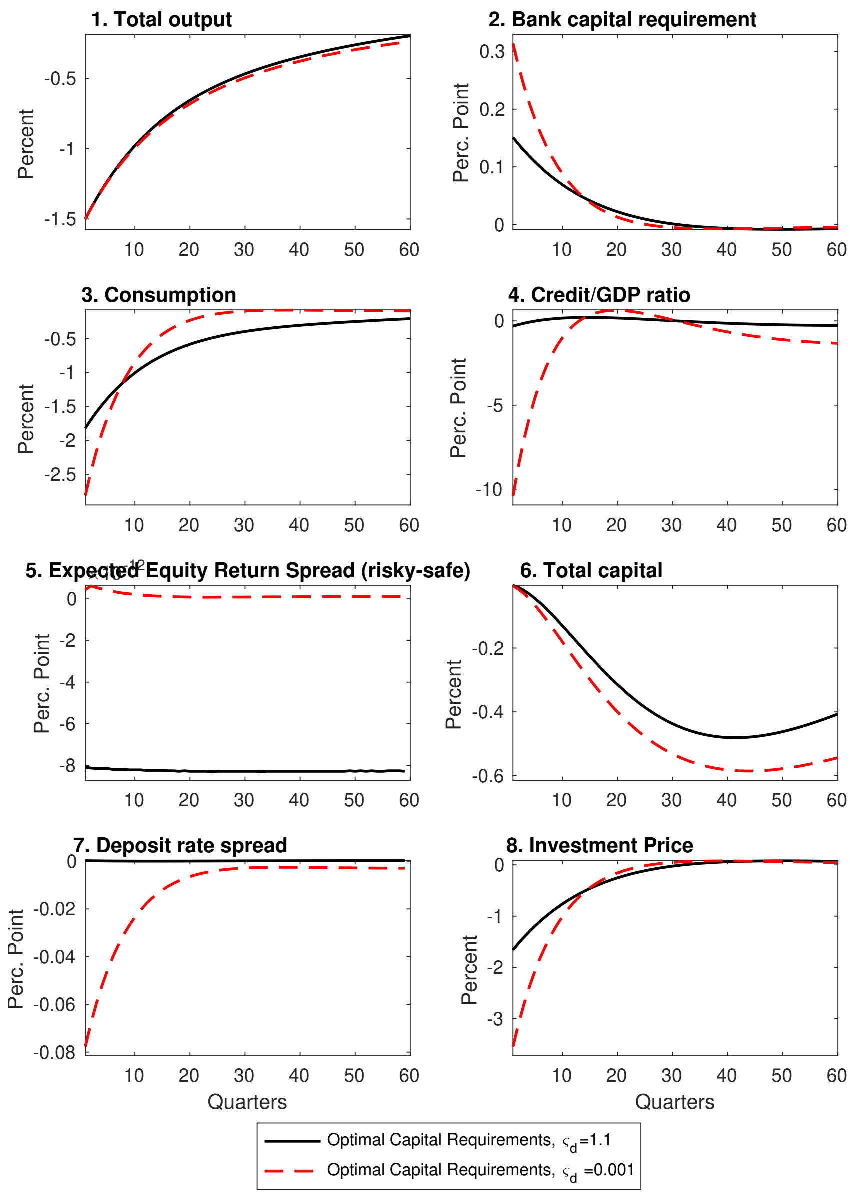


Figure A.5: **Duration of Excessive Risk Taking for Alternative Choices of the Parameter Governing the Interest Elasticity for the Households' Supply of Deposits (elasticity** $= \frac{1}{\varsigma_d}$**)**