A Macroeconomic Model with Financial Panics*

Mark Gertler, Nobuhiro Kiyotaki and Andrea Prestipino
NYU, Princeton and Federal Reserve Board

September, 2017

Abstract

This paper incorporates banks and banking panics within a conventional macroeconomic framework to analyze the dynamics of a financial crisis of the kind recently experienced. We are particularly interested in characterizing the sudden and discrete nature of the banking panics as well as the circumstances that makes an economy vulnerable to such panics in some instances but not in others. Having a conventional macroeconomic model allows us to study the channels by which the crisis affects real activity and the effects of policies in containing crises.

*Preliminary. We thank the National Science Foundation and the Macro Financial Modeling group at the University of Chicago for financial support.

1 Introduction

As both Bernanke (2010) and Gorton (2010) argue, at the heart of the recent financial crisis was a series of bank runs that culminated in the precipitous demise of a number of major financial institutions. During the period where the panics were most intense in October 2008, all the major investment banks effectively failed, the commercial paper market froze, and the Reserve Primary Fund (a major money market fund) experienced a run. The distress quickly spilled over to the real sector. Credit spreads rose to Great Depression era levels. There was an immediate sharp contraction in economic activity: From 2008:Q4 through 2009:Q1 real output dropped at an eight percent annual rate, driven mainly by a nearly forty percent drop in investment spending. Also relevant is that this sudden discrete contraction in financial and real economic activity occurred in the absence of any apparent large exogenous disturbance to the economy.

In this paper we incorporate banks and banking panics within a conventional macroeconomic framework - a New Keynesian model with capital accumulation. Our goal is to develop a model where it is possible to analyze both qualitatively and quantitatively the dynamics of a financial crisis of the kind recently experienced. We are particularly interested in characterizing the sudden and discrete nature of banking panics as well as the circumstances that make the economy vulnerable to such panics in some instances but not in others. Having a conventional macroeconomic model allows us to study the channels by which the crisis affects aggregate economic activity and the effects of various policies in containing crises.

Our paper fits into a lengthy literature aimed at adapting core macroeconomic models to account for financial crises[1]. Much of this literature emphasizes the role of balance sheets in constraining borrower from spending when financial markets are imperfect. Because balance sheets tend to strengthen in booms and weaken in recessions, financial conditions work to amplify fluctuations in real activity. Many authors have stressed that this kind of balance sheet mechanism played a central role in the crisis, particularly for banks and households, but also at the height of the crisis for non-financial firms as well. Nonetheless, as Mendoza (2010), He and Krishnamurthy (2017) and Brunnermeier and Sannikov (2014) have emphasized, these models do not capture the highly nonlinear aspect of the crisis. Although the financial mechanisms

[1] See Gertler and Kiyotaki (2011) and Brunnermeier et. al (2013) for recent surveys.

in these papers tend to amplify the effects of disturbances, they do not easily capture sudden discrete collapses. Nor do they tend to capture the run-like behavior associated with financial panics.

Conversely, beginning with Diamond and Dybvig (1983), there is a large literature on banking panics. An important common theme of this literature is how liquidity mismatch, i.e. partially illiquid long-term assets funded by short-term debt, opens up the possibility of runs. Most of the models in this literature, though, are partial equilibrium and highly stylized (e.g. three periods). They are thus limited for analyzing the interaction between financial and real sectors.

Our paper builds on our earlier work - Gertler and Kiyotaki (GK, 2015) and Gertler, Kiyotaki and Prestipino (GKP, 2016) - which analyzed bank runs in an infinite horizon endowment economy. These papers characterize runs as self-fulfilling rollover crises, following the Calvo (1988) and Cole and Kehoe (2001) models of sovereign debt crises. Both GK and GKP emphasize the complementary nature of balance sheet conditions and bank runs. Balance sheet conditions affect not only borrower access to credit but also whether the banking system is vulnerable to a run. In this way the model is able to capture the discrete highly nonlinear nature of a collapse: When bank balance sheets are strong, negative shocks do not push the financial system to the verge of collapse. When they are weak, the same size shock leads the economy into a crisis zone in which a bank run equilibrium exists.[2] Given that GK and GKP analyze runs in the context of an endowment economy, however, the focus is on the effects of panics on the behavior of asset prices and credit spreads. By extending the analysis to a conventional macroeconomic model, we can explicitly capture the effect of the financial collapse on investment, output and employment. In particular, we proceed to show that a calibrated version of our model is capable of capturing the dynamics of key financial and real variables over the course of the recent crisis.

Also related is important recent work on occasionally binding borrowing constraints as a source of nonlinearity in financial crises such as Mendoza (2010) and He and Krishnamurthy (2017). There, in good times the bor-

[2] Some recent examples where self-fulfilling financial crises can emerge depending on the state of the economy include Benhabib and Wang (2013), Bocola and Lorenzoni (2017) and Farhi and Maggiori (2017), and Perri and Quadrini (forthcoming). For further attempts to incorporate bank runs in macro models, see Angeloni and Faia (2013), Cooper and Ross (1998), Martin, Skeie and Von Thadden (2014), Robatto (2014) and Uhlig (2010) for example.

rowing constraint is not binding and the economy behaves much the way it does with frictionless financial markets. However, a negative disturbance can move the economy into a region where the constraint is binding, amplifying the effect of the shock on the downturn. In a similar spirit, Brunnermeier and Sannikov (2014) generate nonlinear dynamics based on the precautionary saving behavior by intermediaries worried about survival in the face of sequence of negative aggregate shocks. Our approach also allows for occasionally binding financial constraints and precautionary saving. However, in quantitative terms, bank runs provide the major source of nonlinearity.

Section 2 presents the behavior of bankers and workers, the sectors where the novel features of the model are introduced. Section 3 describes the features that are standard in the New Keynesian model: the behavior of firms, price setting, investment and monetary policy. Section 4 describes the calibration and presents a variety of numerical exercises designed to illustrate the main features of the model. We conclude the section with an illustration of how the model can capture the dynamics of some of the main features of the recent financial crisis.

2 Model: outline, households, and bankers

The baseline framework is a standard New Keynesian model with capital accumulation. In contrast to the conventional model, each household consists of bankers and workers. Bankers specialize in making loans and thus intermediate funds between households and productive capital. Households may also make these loans directly, but they are less efficient in doing so than bankers.[3] On the other hand, bankers may be constrained in their ability to raise external funds and also may be subject to runs. The net effect is that the cost of capital will depend on the endogenously determined flow of funds between intermediated and direct finance.

We distinguish between capital at the beginning of period t, K_t, and capital at the end of the period, S_t. Capital at the beginning of the period is used in conjunction with labor to produce output at t. Capital at the end of period is the sum of newly produced capital and the amount of capital left

[3] As section 2.2. makes clear, technically it is the workers within the household that are left to manage any direct finance. But since these workers collectively decide consumption, labor and portfolio choice on of behalf the household, we simply refer to them as the 'household' going forward.

after production:

$$S_t = \Gamma\left(\frac{I_t}{K_t}\right) K_t + (1-\delta)K_t, \tag{1}$$

where δ is the rate of depreciation. The quantity of newly produced capital, $\Gamma(I_t/K_t)K_t$, depends upon investment I_t and the capital stock. We suppose that $\Gamma(\cdot)$ is an increasing and concave function of I_t/K_t to capture convex adjustment costs.

A firm wishing to finance new investment as well as old capital issues a state-contingent claim on the earnings generated by the capital. Let S_t be the total number of claims (effectively equity) outstanding at the end of period t (one claim per unit of capital), S_t^b be the quantity intermediated by bankers and S_t^h be the quantity directly held by households. Then we have:

$$S_t^b + S_t^h = S_t. \tag{2}$$

Both the total capital stock and the composition of financing are determined in equilibrium.

The capital stock entering the next period K_{t+1} differs from S_t due to a multiplicative "capital quality" shock, ξ_{t+1}, that randomly transforms the units of capital available at $t+1$.

$$K_{t+1} = \xi_{t+1} S_t. \tag{3}$$

The shock ξ_{t+1} provides an exogenous source of variation in the return to capital.

To capture that households are less efficient than bankers in handling investments, we assume that they suffer a management cost that depends on the share of capital they hold, S_t^h/S_t. The management cost reflects their disadvantage relative to bankers in evaluating and monitoring investment projects. The cost is in utility terms and takes the following piece-wise form:

$$\varsigma(S_t^h, S_t) = \begin{cases} \frac{\chi}{2}\left(\frac{S_t^h}{S_t} - \gamma\right)^2 S_t, \text{ if } \frac{S_t^h}{S_t} > \gamma > 0 \\ 0, \text{ otherwise} \end{cases} \tag{4}$$

with $\chi > 0$.[4]

[4] For a deeper model of the costs that non experts face in financial markets see Kurlat (2016).

For $S_t^h/S_t \leq \gamma$ there is no efficiency cost: Households are able to manage a limited fraction of capital as well as bankers. As the share of direct finance exceeds γ, the efficiency cost $\varsigma(\cdot)$ is increasing and convex in S_t^h/S_t. In this region, constraints on the household's ability to manage capital become relevant. The convex form implies that the marginal efficiency losses rise with the size of the household's direct capital holdings, capturing limits on its capacity to handle investments.

We assume that the efficiency cost is homogenous in S_t^h and S_t to simplify the computation. As the marginal efficiency cost is linear in the share S_t^h/S_t, it reduces the nonlinearity in the model. An informal motivation is that, as the capital stock S_t increases, the household has more options from which to select investments that it is better able to manage, which works to dampen the marginal efficiency cost.

Given the efficiency costs of direct household finance, absent financial frictions banks will intermediate at least the fraction $1-\gamma$ of the capital stock. However, when banks are constrained in their ability to obtain external funds, households will directly hold more than the share γ of the capital stock. As the constraints tighten in a recession, as will happen in our model, the share of capital held by households will expand. As we will show, in general equilibrium, the reallocation of capital holdings from banks to less efficient households raises the cost of capital, reducing investment and output. In the extreme event of a systemic bank run, the contraction will become far more severe: As banks liquidate all their holdings, the worker share of finance will temporarily rise to unity. In turn, the firesale of assets from banks to inefficient households will lead to a sharp rise in the cost of credit, leading to an extreme contraction in investment and output.

In the rest of this section we characterize the behavior of households and bankers which are the non-standard parts of the model.

2.1 Households

We formulate this sector in a way that allows for financial intermediation yet preserves the tractability of the representative household setup. In particular, each household (family) consists of a continuum of members with measure unity. Within the household there are $1-f$ workers and f bankers. Workers supply labor and earn wages for the household. Each banker manages a bank and transfers non-negative dividend back to the household. Within the family there is perfect consumption insurance.

In order to preclude a banker from retaining sufficient earnings to permanently relax any financial constraint, we assume the following: In each period, with i.i.d. probability $1-\sigma$, a banker exits. Upon exit it then gives all its accumulated earnings to the household. This stochastic exit in conjunction with the payment to the household upon exit is in effect a simple way to model dividend payouts.[5]

After exiting, a banker returns to being a worker. To keep the population of each occupation constant, each period, $(1-\sigma) f$ workers become bankers. At this time the household provides each new banker with an exogenously given initial equity stake in the form of a wealth transfer, e_t. The banker receives no further transfers from the household and instead operates at arms length.

Households save in the form of direct claims on capital and deposits at banks. Bank deposits at t are one period bonds that promise to pay a non-contingent gross real rate of return $\overline{R}_{t+1}$ in the absence of default. In the event of default at $t+1$, depositors receive the fraction x_{t+1} of the promised return, where the recovery rate $x_{t+1} \in [0,1)$ is the total liquidation value of bank assets per unit of promised deposit obligations.

There are two reasons the bank may default: First, a sufficiently negative return on its portfolio may make it insolvent. Second, even if the bank is solvent at normal market prices, the bank's creditors may "run" forcing the bank to liquidate assets at firesale prices. We describe each of these possibilities in detail in the next section. Let p_t be the probability that the bank defaults in period $t+1$. Given p_t and x_t, we can express the gross rate of return on the deposit contract R_{t+1} as

$$R_{t+1} = \begin{cases} \overline{R}_{t+1} \text{ with probability } 1-p_t \\ x_{t+1}\overline{R}_{t+1} \text{ with probability } p_t \end{cases}.$$

Similar to the Cole and Kehoe (2001) model of sovereign default, a run in our model will correspond to a panic failure of households to roll over deposits. This contrasts with the "early withdrawal" mechanism in the classic Diamond and Dybvig (1983) model. For this reason we do not need to impose a "sequential service constraint" which is necessary to generate runs in Diamond and Dybvig. Instead we make the weaker assumption that all households receive the same pro rata share of output in the event of default,

[5] As section 2.2 makes clear, because of the financial constraint, it will always be optimal for a bank to retain earnings until exit.

whether it be due to insolvency or a run. Later we describe the conditions that lead to the existence of an equilibrium where a "failure to rollover" run is possible.

Let C_t be consumption, L_t labor supply, and $\beta \in (0,1)$ the household's subjective discount factor. As mentioned before, $\varsigma(S_t^h, S_t)$ is the household utility cost of direct capital holding S_t^h, where the household takes the aggregate quantity of claims S_t as given. Then household utility U_t is given by

$$U_t = E_t \left\{ \sum_{\tau=t}^{\infty} \beta^{\tau-t} \left[\frac{(C_\tau)^{1-\gamma_h}}{1-\gamma_h} - \frac{(L_\tau)^{1+\varphi}}{1+\varphi} - \varsigma(S_\tau^h, S_\tau) \right] \right\},$$

Let Q_t be the relative price of capital, Z_t the rental rate on capital, w_t the real wage, T_t lump sum taxes, and Π_t dividend distributions net transfers to new bankers, all of which the household takes as given. Then the household chooses C_t, L_t S_t^h and deposits D_t to maximize expected utility subject to the budget constraint

$$C_t + D_t + Q_t S_t^h = w_t L_t - T_t + \Pi_t + R_t D_{t-1} + \xi_t [Z_t + (1-\delta) Q_t] S_{t-1}^h. \quad (5)$$

The first order condition for labor supply is given by:

$$w_t \lambda_t = (L_t)^{\varphi}, \quad (6)$$

where $\lambda_t \equiv (C_t)^{-\gamma_h}$ denotes the marginal utility of consumption.

The first order condition for bank deposits takes into account the possibility of default and is given by

$$1 = [(1-p_t) E_t (\Lambda_{t+1} | no\ def) + p_t E_t (\Lambda_{t+1} x_{t+1} | def)] \cdot \overline{R}_{t+1} \quad (7)$$

where $E_t(\cdot \mid no\ def)$ (and $E_t(\cdot \mid def)$) are expected value of $\cdot$ conditional on no default (and default) at date t+1. The stochastic discount factor Λ_{t+1} satisfies

$$\Lambda_{t+1} = \beta \frac{\lambda_{t+1}}{\lambda_t}. \quad (8)$$

Observe that the promised deposit rate $\overline{R}_{t+1}$ that satisfies equation (7) depends on the default probability p_t as well as the recovery rate x_{t+1}.[6]

[6] Notice that we are already using the fact that in equilibrium all banks will choose the same leverage so that all deposits have the same probability of default.

Finally, the first order condition for capital holdings is given by

$$E_t\left[\Lambda_{t+1}\xi_{t+1}\frac{Z_{t+1}+(1-\delta)Q_{t+1}}{Q_t+\frac{\partial\varsigma(S_t^h,S_t)}{\partial S_t^h}/\lambda_t}\right]=1, \tag{9}$$

where

$$\frac{\partial\varsigma(S_t^h,S_t)}{\partial S_t^h}/\lambda_t = Max\left[\chi\left(\frac{S_t^h}{S_t}-\gamma\right)/\lambda_t,\ 0\right] \tag{10}$$

is the household's marginal cost of direct capital holding.

The first order condition given by (9) will be key in determining the market price of capital. Observe that the market price of capital will tend to be decreasing in the share of capital held by households above the threshold γ since the efficiency cost $\varsigma(S_t^h, S_t)$ is increasing and convex. As will become clear, in a panic run banks will sell all their securities to households, leading to a sharp contraction in asset prices. The severity of the drop will depend on the curvature of the efficiency cost function given by (4), which controls asset market liquidity in the model.

2.2 Bankers

The banking sector we characterize corresponds best to the shadow banking system which was at the epicenter of the financial instability during the Great Recession. In particular, banks in the model are completely unregulated, hold long-term securities, issue short-term debt, and as a consequence are potentially subject to runs.

2.2.1 Bankers optimization problem

Each banker manages a financial intermediary with the objective of maximizing the expected utility of the household. Bankers fund capital investments by issuing short term deposits d_t to households as well as by using their own equity, or net worth, n_t. Due to financial market frictions, described later, bankers may be constrained in their ability to obtain deposits.

So long as there is a positive probability that the banker may be financially constrained at some point in the future, it will be optimal for the banker to delay dividend payments until exit (as we will verify later). At this point the dividend payout will simply be the accumulated net worth. Accordingly, we can take the banker's objective as to maximize the discounted expected value of net worth upon exit. Given that σ is the survival probability and given that the banker uses the household's intertemporal marginal rate of substitution $\widetilde{\Lambda}_{t,\tau} = \beta^{\tau-t}\lambda_\tau/\lambda_t$ to discount future payouts, we can express the objective of a continuing banker at the end of period t as

$$\begin{aligned} V_t &= E_t\left[\sum_{\tau=t+1}^{\infty} \widetilde{\Lambda}_{t,\tau}(1-\sigma)\sigma^{\tau-t-1}n_\tau\right] \\ &= E_t\left\{\Lambda_{t+1}[(1-\sigma)n_{t+1} + \sigma V_{t+1}]\right\}, \end{aligned} \tag{11}$$

where $(1-\sigma)\sigma^{\tau-t-1}$ is probability of exiting at date τ, and n_τ is terminal net worth if the banker exits at τ.

During each period t, a continuing bank (either new or surviving) finances asset holdings $Q_t s_t^b$ with newly issued deposits and net worth:

$$Q_t s_t^b = d_t + n_t. \tag{12}$$

We assume that banks can only accumulate net worth by retained earnings and do not issue new equity. While this assumption is a reasonable approximation of reality, we do not explicitly model the agency frictions that underpin it.

The net worth of "surviving" bankers, accordingly, is the gross return on assets net the cost of deposits, as follows:

$$n_t = R_t^b Q_{t-1} s_{t-1}^b - R_t d_{t-1}, \tag{13}$$

where R_t^b is the gross rate of return on capital intermediated by banks as

$$R_t^b = \xi_t \frac{Z_t + (1-\delta)Q_t}{Q_{t-1}}. \tag{14}$$

So long as n_t is strictly positive the bank does not default. In this instance it pays its creditors the promised rate $\overline{R}_t$. If the value of assets, $R_t^b Q_{t-1} s_{t-1}^b$, is below the promised repayments to depositors $\bar{R}_t d_{t-1}$ (due to either a run or simply a bad realization of returns), n_t goes to zero and the bank defaults. It then pays creditors the product of recovery rate x_t and $\overline{R}_t$, where x_t is given by.

$$x_t = \frac{R_t^b Q_{t-1} s_{t-1}^b}{\overline{R}_t d_{t-1}} < 1. \tag{15}$$

For new bankers at t, net worth simply equals the start-up equity e_t it receives from the household.

$$n_t = e_t. \tag{16}$$

To motivate a limit on a bank's ability to issue deposits, we introduce the following moral hazard problem: After accepting deposits and buying assets at the beginning of t, but still during the period, the banker decides whether to operate "honestly" or to divert assets for personal use. To operate honestly means holding assets until the payoffs are realized in period $t+1$ and then meeting deposit obligations. To divert means selling a fraction θ of assets secretly on a secondary market in order to obtain funds for personal use. We assume that the process of diverting assets takes time: The banker cannot quickly liquidate a large amount of assets without the transaction being noticed. Accordingly, the banker must decide whether to divert at t, prior to the realization of uncertainty at $t+1$. Further, to remain undetected, he can only sell up to a fraction θ of the assets. The cost to the banker of the diversion is that the depositors force the intermediary into bankruptcy at the beginning of the next period.[7]

The banker's decision on whether or not to divert funds at t boils down to comparing the franchise value of the bank V_t, which measures the present discounted value of future payouts from operating honestly, with the gain from diverting funds, $\theta Q_t s_t^b$. In this regard, rational depositors will not lend to the banker if he has an incentive to cheat. Accordingly, any financial arrangement between the bank and its depositors must satisfy the incentive constraint:

$$\theta Q_t s_t^b \leq V_t. \tag{17}$$

To characterize the banker's optimization problem it is useful to let ϕ_t denote the bank's ratio of assets to net worth, $Q_t s_t^b / n_t$, which we will call the "leverage multiple." Then, combining the flow of funds constraint (13) and the balance sheet constraint (12) yields the expression for the evolution of net worth for a surviving bank as:

$$n_{t+1} = [(R_{t+1}^b - \overline{R}_{t+1})\phi_t + \overline{R}_{t+1}]n_t. \tag{18}$$

[7]We assume households deposit funds in banks other than they ones they own. Hence, diverting involves stealing funds from families other than the one to which the banker belongs.

Using the evolution of net worth equation (18) in the expression for the franchise value of the bank (11) we can write

$$V_t = (\mu_t \phi_t + \nu_t)\, n_t,$$

where

$$\mu_t = (1 - p_t) E_t\{\Omega_{t+1}(R^b_{t+1} - \overline{R}_{t+1}) \mid no\ def\} \tag{19}$$

$$\nu_t = (1 - p_t) E_t\{\Omega_{t+1}\overline{R}_{t+1} \mid no\ def\} \tag{20}$$

$$\Omega_{t+1} = \Lambda_{t+1}(1 - \sigma + \sigma\psi_{t+1})$$

with

$$\psi_{t+1} \equiv \frac{V_{t+1}}{n_{t+1}}.$$

The variable μ_t is the expected discounted excess return on banks assets relative to deposits and ν_t is the expected discounted cost of a unit of deposits. Intuitively, $\mu_t \phi_t$ is the excess return the bank receives from having on additional unity of net worth (taking into account the ability to increase leverage), while ν_t is the cost saving from substituting equity finance for deposit finance.

Notice that the bank uses the stochastic discount factor Ω_{t+1} to value returns in $t + 1$. Ω_{t+1} is the banker's discounted shadow value of a unit of net worth at $t + 1$, averaged across the likelihood of exit and the likelihood of survival. We can think of ψ_{t+1} in the expression for Ω_{t+1} as the bank's "Tobin's Q ratio", i.e., the ratio of the franchise value to the replacement cost of the bank balance sheet. With probability $1 - \sigma$ the banker exits, implying the discounted shadow value of a unit of net worth simply equals the household discount factor Λ_{t+1}. With probability σ the banker survives implying the discounted marginal value of n_{t+1} equals the discounted value of the bank's Tobin's Q ratio, $\Lambda_{t+1}\psi_{t+1}$. As will become clear, to the extent that an additional unit of net worth relaxes the financial market friction, ψ_{t+1} in general will exceed unity provided that the bank does not default.

The banker's optimization problem is then to choose the leverage multiple ϕ_t to solve

$$\max_{\phi_t} \ (\mu_t \phi_t + \nu_t)\,, \tag{21}$$

subject to the incentive constraint (obtained from equation (17)):

$$\theta\phi_t \leq \mu_t \phi_t + \nu_t, \tag{22}$$

and the deposit rate constraint (obtained from equations (7) and (15)):

$$\overline{R}_{t+1} = \left[(1-p_t)E_t(\Lambda_{t+1} \mid no\ def) + p_t E_t(\Lambda_{t+1}x_{t+1} \mid def)\right]^{-1}, \tag{23}$$

where x_{t+1} is the following function of ϕ_t:

$$x_{t+1} = \frac{\phi_t}{\phi_t - 1}\frac{R^b_{t+1}}{\overline{R}_{t+1}}.$$

and μ_t and ν_t are given by $(19)-(20)$. Notice that since individual bank net worth does not appear in the bank optimization problem, the optimal choice of ϕ_t is independent of n_t.

2.2.2 Banker's decision rules

Let μ^r_t be the expected discounted *marginal return to* increasing the leverage multiple[8]

$$\mu^r_t = \frac{d\psi_t}{d\phi_t} = \mu_t - (\phi_t - 1)\frac{\nu_t}{\overline{R}_{t+1}}\frac{d\overline{R}_{t+1}(\phi_t)}{d\phi_t} < \mu_t. \tag{24}$$

The second term on the right of equation (24) reflects the effect of the increase in $\overline{R}_{t+1}$ that arises as the bank increases ϕ_t. An increase in ϕ_t reduces the recovery rate, forcing $\overline{R}_{t+1}$ up to compensate depositors, as equation (23) suggests. The term $(\phi_t - 1)\,\nu_t/\overline{R}_{t+1}$ then reflects the reduction in the bank franchise value that results from a unit increase in $\overline{R}_{t+1}$. Due to the effect on $\overline{R}_{t+1}$ from expanding ϕ_t, the marginal return μ^r_t is below the average excess return μ_t.

The solution for ϕ_t depends on whether or not the incentive constraint (22) is binding. In the case where (22) binds, making use of (22) implies the following solution for ϕ_t:

$$\phi_t = \frac{\nu_t}{\theta - \mu_t}, \quad \text{if } \mu^r_t > 0. \tag{25}$$

[8] Note that, although the default probability p_t depends upon ϕ_t, the marginal effect of ϕ_t on firm value V_t through the change of p_t is zero. This is because at the borderline of default, $n_{t+1} = 0$ and thus $V_{t+1} = 0$. Thus a small shift in the probability mass from the no-default to the default region has no impact on V_t. Similarly, the promised deposit rate $\overline{R}_t$ does not change since at the borderline of default, the recovery rate x_t is unity. See Appendix for details. Important to the argument is the absence of deadweight loss associated with default.

In this instance, even though the marginal return to increasing the leverage multiple is positive, the incentive constraint limits the bank from increasing leverage to acquire more assets. The constraint (25) limits the leverage multiple to the point where the bank's gain from diverting funds per unit of net worth $\theta\phi_t$ is exactly balanced by the cost per unit of net worth of losing the franchise value, which is measured by $\psi_t = \mu_t\phi_t + \nu_t$. Note that μ_t tends to move countercyclically since the excess return on bank capital $E_t R^b_{t+1} - \overline{R}_{t+1}$ widens as the borrowing constraint tightens in recessions. As a result, ϕ_t tends to move countercyclically. As we show later, the countercyclical movement in ϕ_t contributes to making bank runs more likely in bad economic times.[9]

Conversely, when the constraint is not binding now, the bank expands leverage and assets to the point where the marginal return to increasing the leverage multiple is zero as,

$$\mu^r_t = 0, \quad \text{if } \phi_t < \frac{\nu_t}{\theta - \mu_t}. \tag{26}$$

Even if the constraint does not bind, the bank may still choose to limit the leverage multiple, so long as there is a possibility that the incentive constraint could bind in the future. In this instance, as in Brunnermeier and Sannikov (2014) and He and Krishnamurthy (2015), banks have a precautionary motive for scaling back their respective leverage multiples.[10] The precautionary motive is reflected by the presence of the discount factor Ω_{t+1} in the measure of the discounted excess return. The discount factor Ω_{t+1}, which reflects the

[9] In the data, net worth of our model corresponds to the mark-to-market difference between assets and liabilities of the bank balance sheet. It is different from the book value often used in the official report, which is slow in reacting to market conditions. Also bank assets here are securities and loans to the non-financial sector, which exclude those to other financial intermediaries. In the data, the net mark-to-market leverage multiple of the financial intermediation sector - the ratio of securities and loans to the nonfinancial sector to the net worth of the aggregate financial intermediaries - tends to move counter-cyclically, even though the gross leverage multiple - the ratio of book value total assets (including securities and loans to the other intermediaries) to the net worth of some individual intermediaries may move procyclically. Concerning the debate about the procyclicality and countercyclicality of the leverage rate of the intermediaries, see Adrian and Shin (2010) and He, Khang and Krishnamurthy (2010).

[10] One difference from these papers is that because default is possible, the bank's decision over its leverage multiple also affects the promised deposit rate, which affects the cost of funds at the margin. This effect provides an additional motive for the bank to reduce its leverage multiple.

shadow value of net worth, tends to vary countercyclically given that borrowing constraints tighten in downturns. By reducing their leverage multiples, banks reduce the risk of taking losses when the shadow value of net worth is high.

In either case, as we conjectured, the franchise value of the bank V_t is proportionate to n_t by a factor that is independent of bank-specific factors: When the incentive constraint is binding:

$$V_t = \theta\phi_t \cdot n_t$$

as equation (22) suggests. When it is not currently binding,

$$V_t = \left\{ \left[(\phi_t - 1) \frac{\nu_t}{\overline{R}_{t+1}} \frac{d\overline{R}_{t+1}(\phi_t)}{d\phi_t} \right] \phi_t + \nu_t \right\} \cdot n_t$$

as equations (21), (24) and (26) suggest.

An important corollary is that the bank cannot operate with zero net worth. In this instance V_t falls to zero, implying that the incentive constraint (17) would always be violated if the bank tried to issue deposits. That banks require positive equity to operate is vital to the possibility of bank runs. As we show, a necessary condition for a bank run equilibrium to exist is that banks cannot operate with zero net worth.

2.2.3 Aggregation of the financial sector absent default

We now characterize the aggregate financial sector during periods where banks do not default. We then turn to the case of default due either to runs or insolvency.

Given that individual bank portfolio decisions are homogenous in net worth, the optimal leverage multiple ϕ_t is independent of bank-specific factors. Accordingly, we can sum across banks to obtain the following relation between aggregate bank asset holdings $Q_t K_t^b$ and the aggregate quantity of net worth N_t in the banking sector:

$$\frac{Q_t K_t^b}{N_t} = \phi_t. \tag{27}$$

We next characterize the evolution of N_t which depends on both the retained earnings of bankers that survived from the previous period and the injection of equity to new bankers. For technical convenience again related to computational considerations, we suppose that the household transfer e_t to a each new banker is proportionate to the stock of capital at the end of the previous period, S_{t-1}, with $e_t = \frac{\zeta}{(1-\sigma)f} S_{t-1}$.[11] Aggregating across both surviving and entering bankers yields the following expression for the evolution of net worth

$$N_t = \sigma[(R_t^b - \overline{R}_t)\phi_{t-1} + \overline{R}_t]N_{t-1} + \zeta S_{t-1}. \tag{28}$$

The first term is the total net worth of bankers that operated at $t-1$ and survived until t. The second, ζS_{t-1}, is the total start-up equity of entering bankers.

2.3 Runs versus insolvency and the default probability

In this section we describe bank runs and the condition for a bank run equilibrium to exist. We distinguish a run equilibrium due to illiquidity from insolvency. We then characterize the overall default probability. Within our calibrated model, the probability of runs will significantly increase the likelihood of default.

2.3.1 Conditions for a bank run equilibrium

As in Diamond and Dybvig (1983), the runs we consider are runs on the entire banking system and not an individual bank. A run on an individual bank will not have aggregate effects as depositors simply shuffle their funds from one bank to another. As we noted earlier, though, we differ from Diamond and Dybvig in that runs reflect a panic failure to roll over deposits as opposed to early withdrawal.

Consider the behavior of a household that acquired deposits at $t-1$. Suppose further that the banking system is solvent at the beginning of time t : Net worth is positive, implying that assets valued at normal market prices exceed liabilities. The household must then decide whether to roll over deposits at t. A self-fulfilling "run" equilibrium exists if the household perceives that in the event all other depositors run, thus forcing the banking system into liquidation, the household will lose money if it rolls over its deposits individually. Note that this condition is satisfied if the liquidation makes the banking system insolvent, i.e. drives aggregate bank net worth to

[11] Here we value capital at the steady state price $Q = 1$. If we use the market price instead, the financial accelerator would be enhanced but not significantly.

zero. A household that deposits funds in a zero net worth bank will simply lose its money as the bank will divert the money for personal use.

The condition for a bank run equilibrium at t, accordingly, is that in the event of liquidation following a run, bank net worth goes to zero. Recall that earlier we defined the depositor recovery rate, x_t, as the ratio of the value of bank assets in liquidation to promised obligations to depositors. Accordingly, the condition for a bank run equilibrium is simply that the recovery rate conditional on a run, x_t^R, is less than unity:

$$\begin{aligned} x_t^R &= \frac{\xi_t[(1-\delta)Q_t^* + Z_t^*]S_{t-1}^b}{\overline{R}_t D_{t-1}} \\ &= \frac{R_t^{b*}}{\overline{R}_t} \cdot \frac{\phi_{t-1}}{\phi_{t-1}-1} < 1 \end{aligned} \tag{29}$$

where Q_t^* is the asset liquidation price, Z_t^* is rental rate, and R_t^{b*} is the return on bank assets conditional on run. Note that in general the liquidation price Q_t^* is below the normal market price Q_t, implying that a run may occur even if the bank is solvent at normal market prices. Further, as will be shown later, given $\frac{R_t^{b*}}{\overline{R}_t}$ is procyclical and ϕ_{t-1} is countercyclical, the likelihood of a bank run equilibrium existing is greater in recessions than in booms.

2.3.2 The liquidation price

Key to the condition for a bank run equilibrium is the behavior of the liquidation price Q_t^*. A depositor run at t induces all the existing banks to liquidate their assets by selling them to households. We suppose that new banks can only store their net worth during a run and start raising deposit one period after the panic. Accordingly in the wake of the run:

$$S_t^h = S_t. \tag{30}$$

The banking system then rebuilds itself over time as new banks enter. The evolution of net worth following the run at t is given by

$$\begin{aligned} N_{t+1} &= \zeta S_t \left(1 + \sigma \frac{S_{t-1}}{S_t}\right). \\ N_\tau &= \sigma[(R_\tau^b - R_\tau)\phi_{\tau-1} + R_\tau]N_{\tau-1} + \zeta S_{\tau-1}, \text{ for all } \tau \geq t+2. \end{aligned} \tag{31}$$

To obtain Q_t^*, we invert the household Euler equation to obtain:

$$
\begin{aligned}
Q_t^* &= E_t\left\{\sum_{\tau=t+1}^{\infty}\widetilde{\Lambda}_{t,\tau}(1-\delta)^{\tau-t-1}\left(\prod_{j=t+1}^{\tau}\xi_j\right)\cdot\left[Z_\tau-\chi\left(\frac{S_\tau^h}{S_\tau}-\gamma\right)/\lambda_\tau\right]\right\} \\
&\quad -\chi\left(1-\gamma\right)/\lambda_t. \qquad (32)
\end{aligned}
$$

where the term $\chi\left(1-\gamma\right)/\lambda_t$ is the period t marginal efficiency cost following a run at t (given $S_t^h/S_t=1$ in this instance) [12] The liquidation price is thus equal to the expected discounted stream of dividends net the marginal efficiency losses from household portfolio management. Since marginal efficiency losses are at a maximum when S_t^h equal S_t, Q_t^* is at a minimum, given the expected future path of S_t^h. Further, the longer it takes the banking system to recover (so S_t^h falls back to its steady state value) the lower will be Q_t^*. Finally, note that Q_t^* will vary positively with the expected path of ξ_τ and Z_τ and with the stochastic discount factor $\widetilde{\Lambda}_{t,\tau}$.

2.3.3 The default probability and illiquidity versus insolvency

In the run equilibrium, banks default even though they are solvent at normal market prices. It is the forced liquidation at firesale prices during a run that pushes these banks into bankruptcy. Thus, in the context of our model, a bank run can be viewed as a situation of illiquidity. By contrast, default is also possible if banks enter period t insolvent at normal market prices.

Accordingly, the total probability of default in the subsequent period, p_t, is the sum of the probability of a run p_t^R and the probability of insolvency p_t^I :

$$p_t = p_t^R + p_t^I. \qquad (33)$$

We begin with p_t^I. By definition, banks are insolvent if the ratio of assets valued at normal market prices is less than liabilities. In our economy, the only exogenous shock to the aggregate economy is a shock to quality of capital ξ_t. Accordingly, define ξ_{t+1}^I as the value of capital quality, ξ_{t+1}, that makes the depositor recovery rate at normal market prices, $x(\xi_{t+1}^I)$ equal to unity.

$$x(\xi_{t+1}^I) = \frac{\xi_{t+1}^I[Z_{t+1}(\xi_{t+1}^I)+(1-\delta)Q_{t+1}(\xi_{t+1}^I)]S_t^b}{\overline{R}_t D_t} = 1. \qquad (34)$$

[12] We are imposing that $\frac{S_t^h}{S_t}-\gamma\geq 0$ as is the case in all of our numerical simulations.

For values of ξ_{t+1} below ξ^I_{t+1}, the bank will be insolvent and must default. Accordingly, the probability of default due to insolvency is given by

$$p^I_t = prob_t\left(\xi_{t+1} < \xi^I_{t+1}\right), \tag{35}$$

where $prob_t(\cdot)$ is the probability of satisfying $\cdot$ conditional on date t information.

We next turn to the determination of the run probability. In general, the time t probability of a run at $t+1$ is product of the probability a run equilibrium exists at $t+1$ times the probability a run will occur when it's feasible. We suppose the latter depends on the realization of a sunspot. Let ι_{t+1} be a binary sunspot variable that takes on a value of 1 with probability $\varkappa$ and a probability of 0 with probability $1-\varkappa$. In the event of $\iota_{t+1} = 1$, depositors coordinate on a run if a bank run equilibrium exists. Note that we make the sunspot probability $\varkappa$ constant so as not to build in exogenous cyclicality in the movement of the overall bank run probability p^R_t.

Accordingly, a bank run arises at $t+1$ iff (i) a bank run equilibrium exists at $t+1$ and (ii) $\iota_{t+1} = 1$. Let ω_t be the probability at t that a bank run equilibrium exists at $t+1$. Then the probability p^R_t of a run at $t+1$ is given by

$$p^R_t = \omega_t \cdot \varkappa. \tag{36}$$

To find the value of ω_t, let us define ξ^R_{t+1} as the value of ξ_{t+1} that makes the recovery rate conditional on a run x^R_{t+1} unity when evaluated at the firesale liquidation price Q^*_{t+1}:

$$x(\xi^R_{t+1}) = \frac{\xi^R_{t+1}[(1-\delta)Q^*(\xi^R_{t+1}) + Z(\xi^R_{t+1})]S^b_t}{\overline{R}_t D_t} = 1. \tag{37}$$

Accordingly, for values of ξ_{t+1} below ξ^R_{t+1}, x^R_{t+1} is below unity, a bank run equilibrium is feasible. The probability of a bank run equilibrium existing is accordingly the probability that ξ_{t+1} lies in the interval below ξ^R_{t+1} but above the threshold for insolvency ξ^I_{t+1}. In particular,

$$\omega_t = prob_t\left(\xi^I_{t+1} \leq \xi_{t+1} < \xi^R_{t+1}\right). \tag{38}$$

Given equation (38), we can distinguish regions of ξ_{t+1} where insolvency emerges ($\xi_{t+1} < \xi^I_{t+1}$) from regions where an illiquidity problem may emerge ($\xi^I_{t+1} \leq \xi_{t+1} < \xi^R_{t+1}$).

Overall, the probability of a run varies inversely with the expected recovery rate $E_t x_{t+1}$. The lower the forecast of the depositor recovery rate, the higher ω_t and thus the higher p_t. In this way the model captures that an expected weakening of the banking system raises the likelihood of a run.

Finally, comparing equations (35) and (38) makes clear that the possibility of a run equilibrium expands the set of realizations where default is possible. That is, the possibility of runs significantly expands the chances for a banking collapse, beyond the probability that would arise simply from default due to insolvency. In this way the possibility of runs makes the system more fragile. Indeed, within the numerical exercises we present the probability of a fundamental shock that induces an insolvent banking system is negligible. However, the probability of a shock that induces a bank run equilibrium is non-trivial.

3 Production, market clearing and policy

The rest of the model is fairly standard. There is a production sector consisting of producers of final goods, intermediate goods and capital goods. Prices are sticky in the intermediate goods sector. In addition there is a central bank that conducts monetary policy.

3.1 Final and intermediate goods firms

There is a continuum of measure unity of final goods producers and intermediate goods producers. Final goods firms make a homogenous good Y_t that may be consumed or used as input to produce new capital goods. Each intermediate goods firm $f \in [0, 1]$ makes a specialize good $Y_t(f)$ that is used in the production of final goods.

The production function that final goods firms use to transforms intermediate goods into final output is given by the following CES aggregator:

$$Y_t = \left[\int_0^1 Y_t\left(f\right)^{\frac{\varepsilon-1}{\varepsilon}} df \right]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{39}$$

where $\varepsilon > 1$ is the elasticity of substitution between intermediate goods.

Let $P_t\left(f\right)$ be the nominal price of intermediate good f. Then cost minimization yields the following demand function for each intermediate good f

(after integrating across the demands of by all final goods firms):

$$Y_t(f) = \left[\frac{P_t(f)}{P_t}\right]^{-\varepsilon} Y_t, \tag{40}$$

where P_t is the price index as

$$P_t = \left[\int_0^1 P_t(f)^{1-\varepsilon} df\right]^{\frac{1}{1-\varepsilon}}.$$

There is a continuum of intermediate good firms owned by consumers, indexed by $f \in [0, 1]$. Each produces a differentiated good and is a monopolistic competitor. Intermediate goods firm f uses both labor $L_t(f)$ and capital $K_t(f)$ to produce output according to:

$$Y_t(f) = A_t K_t(f)^{\alpha} L_t(f)^{1-\alpha}, \tag{41}$$

where A_t is a technology parameter and $0 > \alpha > 1$ is the capital share.

Both labor and capital are freely mobile across firms. Firms rent capital from owners of claims to capital (i.e. banks and households) in a competitive market on a period by period basis. Then from cost minimization, all firms choose the same capital labor ratio, as follows

$$\frac{K_t(f)}{L_t(f)} = \frac{\alpha}{1-\alpha}\frac{w_t}{Z_t} = \frac{K_t}{L_t}. \tag{42}$$

where, as noted earlier, w_t is the real wage and Z_t is the rental rate of capital. The first order conditions from the cost minimization problem imply that marginal cost is given by

$$MC_t = \frac{1}{A_t}\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha}\left(\frac{Z_t}{\alpha}\right)^{\alpha}. \tag{43}$$

Observe that marginal cost is independent of firm-specific factors.

Following Rotemberg (1982), each monopolistically competitive firm f faces quadratic costs of adjusting prices. Let ρ^r ("r" for Rotemberg) be the parameter governing price adjustment costs. Then each period, it chooses $P_t(f)$ and $Y_t(f)$ to maximize the expected discounted value of profit:

$$E_t\left\{\sum_{\tau=t}^{\infty} \widetilde{\Lambda}_{t,\tau}\left[\left(\frac{P_\tau(f)}{P_\tau} - MC_\tau\right) Y_\tau(f) - \frac{\rho^r}{2} Y_\tau\left(\frac{P_\tau(f)}{P_{\tau-1}(f)} - 1\right)^2\right]\right\}, \tag{44}$$

subject to the demand curve (40). Here we assume that the adjustment cost is proportional to the aggregate demand Y_t.

Taking the firm's first order condition for price adjustment and imposing symmetry implies the following forward looking Phillip's curve:

$$(\pi_t - 1)\,\pi_t = \frac{\varepsilon}{\rho^r}\left(MC_t - \frac{\varepsilon - 1}{\varepsilon}\right) + E_t\left[\Lambda_{t+i}\frac{Y_{t+1}}{Y_t}\left(\pi_{t+1} - 1\right)\pi_{t+1}\right], \qquad (45)$$

where $\pi_t = \frac{P_t}{P_{t-1}}$ is the realized gross inflation rate at date t.

3.2 Capital goods producers

There is a continuum of measure unity of competitive capital goods firms. Each produces new investment goods that it sells at the competitive market price Q_t. By investing $I_t(j)$ units of final goods output, firm j can produce $\Gamma(I_t(j)/K_t) \cdot K_t$ new capital goods, with $\Gamma' > 0$, $\Gamma'' < 0$, and where K_t is the aggregate capital stock.[13]

The decision problem for capital producer j is accordingly

$$\max_{I_t(j)} Q_t \Gamma\left(\frac{I_t(j)}{K_t}\right) K_t - I_t(j). \qquad (46)$$

Given symmetry for capital producers ($I_t(j) = I_t$), we can express the first order condition as the following "Q" relation for investment:

$$Q_t = \left[\Gamma'\left(\frac{I_t}{K_t}\right)\right]^{-1} \qquad (47)$$

which yields a positive relation between Q_t and investment.

3.2.1 Monetary Policy

Let Θ_t be a measure of cyclical resource utilization, i.e., resource utilization relative to the flexible price equilibrium. Next let $R = \beta^{-1}$ denote the real

[13] For simplicity we are assuming that the aggregate capital stock enters into production function of investment goods as an externality. Alternatively, we could make an assumpiton similar to Lorenzoni and Walentin (2007): Each capital goods producer buys capital after being used to produce intermediated goods and combines the capital with final output goods to produce the total capital stock. One can then obtain a first order condition like (47).

interest rate in the deterministic steady state with zero inflation. We suppose that the central bank sets the nominal rate on the riskless bond R_t^n according to the following Taylor rule:

$$R_t^n = \frac{1}{\beta} (\pi_t)^{\kappa_\pi} (\Theta_t)^{\kappa_y} \tag{48}$$

with $\kappa_\pi > 1$. Note that, if the net nominal rate cannot go below zero, the policy rule would become $R_t^n = \max\left\{\frac{1}{\beta} (\pi_t)^{\kappa_\pi} (\Theta_t)^{\kappa_y}, 1\right\}$.

A standard way to measure Θ_t is to use the ratio of actual output to a hypothetical flexible price equilibrium value of output. Computational considerations lead us to use a measure which similarly captures the cyclical efficiency of resource utilization but is much easier to handle numerically. Specifically, we take as our measure of cyclical resource utilization the ratio of the desired markup, $1 + \mu = \varepsilon/(\varepsilon - 1)$ to the current markup $1 + \mu_t$.[14]

$$\Theta_t = \frac{1+\mu}{1+\mu_t} \tag{49}$$

with

$$1 + \mu_t = MC_t^{-1} = \frac{(1-\alpha)(Y_t/L_t)}{L_t^{\varphi} C_t^{\gamma_h}}. \tag{50}$$

The markup corresponds to the ratio of the marginal product of labor to the marginal rate of substitution between consumption and leisure, which corresponds to the labor market wedge. The inverse markup ratio Θ_t thus isolates the cyclical movement in the efficiency of the labor market, specifically the component that is due to nominal rigidities.

Finally, one period bonds which have a riskless nominal return have zero net supply. (Bank deposits have default risk). Nonetheless we can use the following household Euler equation to price the nominal interest rate of these bonds R_t^n as

$$E_t \left(\Lambda_{t+1} \frac{R_t^n}{\pi_{t+1}} \right) = 1. \tag{51}$$

[14] In the case of consumption goods only, our markup measure of efficiency corresponds exactly to the output gap.

3.2.2 Resource constraints and equilibrium

Total output is divided between consumption, investment, the adjustment cost of nominal prices and a fixed value of government consumption G:

$$Y_t = C_t + I_t + \frac{\rho^r}{2} (\pi_t - 1)^2 Y_t + G. \tag{52}$$

Given a symmetric equilibrium, we can express total output as the following function of aggregate capital and labor:

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha}. \tag{53}$$

Although we consider a limiting case in which supply of government bond and money is zero, government adjusts lump-sum tax to satisfy the budget constraint. Finally, labor market must clear, which implies that the total quantity of labor demanded must equaled the total amount supply by households.

This completes the description of the model. See Appendix for the detail.

4 Numerical exercises

4.1 Calibration

Table 1 lists the choice of parameter values for our model. Overall there are twenty one parameters. Thirteen are conventional as they appear in standard New Keynesian DSGE models. The other eight parameters govern the behavior of the financial sector, and hence are specific to our model.

We begin with the conventional parameters. For the discount rate β, the risk aversion parameter γ_h, the inverse Frisch elasticity φ, the elasticity of substitution between goods ε, the depreciation rate δ and the capital share α we use standard values in the literature. Three additional parameters (η, a, b) involve the investment technology, which we express as follows:

$$\Gamma\left(\frac{I_t}{K_t}\right) = a\left(\frac{I_t}{K_t}\right)^{1-\eta} + b.$$

We set η, which corresponds to the elasticity of the price of capital with respect to investment rate, equal to 0.25, a value in line with panel data estimates. We then choose a and b to hit two targets: first, a ratio of quarterly investment to the capital stock of 2.5% and, second, a value of the price of capital Q equal to unity in the risk-adjusted steady state. We set the value of fixed government expenditure G to 20% of steady state output. Next we choose the cost of price adjustment parameter ρ^{jr} to generate an elasticity of inflation with respect to marginal cost equal to 1 percent, which is roughly in line with the estimates.[15] Finally, we set the feedback parameters in the Taylor rule, κ_π and κ_y to their conventional values of 1.5 and 0.5 respectively.

We now turn to the financial sector parameters. There are six parameters that directly affect the evolution of bank net worth and credit spreads: the banker's survival probability σ; the initial equity injection to entering bankers as a share of capital ζ; the asset diversion parameter θ; the threshold share for costless direct household financing of capital, γ; the parameter governing the convexity of the efficiency cost of direct financing χ; and the probability of observing a sunspot π.

We choose the values of these parameter to hit the following six targets: (i) the average arrival rate of a systemic bank run equals 4 percent annually, corresponding to a frequency of banking panics of once every 25 years, which is in line with the evidence for advanced economies[16]; (ii) the average bank leverage multiple equals 10;[17] (iii) the average excess rate of return on bank assets over deposits equals 2%, based on Philippon (2015); (iv) the average share of bank intermediated assets equals 0.5, which is a reasonable estimate of the share of intermediation performed by investment banks and large commercial banks; (v) and (vi) the increase in excess returns (measured by credit spreads) and the drop in investment following a bank run match the evidence from the recent crisis.

The remaining two parameters determine the serial correlation of the capital quality ρ_ξ and and the standard deviation of the innovations σ_ξ. That is we assume that the capital quality shock obeys the following first order process :

$$\log \xi_{t+1} = \rho_\xi \log \xi_t + \epsilon_{t+1}$$

with $0 < \rho_\xi < 1$ and where ϵ_{t+1} an normally distributed i.i.d. random variable

[15] See, for example, Del Negro, Giannoni and Shorfheide (2015)

[16] See, for example, Bordo et al (2001), Reinhart and Rogoff (2009) and Schularick and Taylor (2012).

[17] We think of the banking sector in our model as including both investment banks and some large commercial banks that operated off balance sheet vehicles without explicit guarantees. Ten is on the high side for commercial banks and on the low side for investment banks. See Gertler Kiyotaki Prestipino (2016).

with mean zero and standard deviation σ_ξ. We choose ρ_ξ and σ_ξ so that the unconditional standard deviations of investment and output that match the ones observed over the 1983Q1-2008Q3 period.

Given that our policy functions are non linear we obtain model implied moments by simulating our economy for 100 thousand periods. Table 2 shows unconditional standard deviations for some key macroeconomic variables in the model and in the data. The volatilities of output, investment and labor are reasonably in line with the data. Consumption is too volatile, but the variability of the aggregate of consumption and investment matches the evidence.

4.2 Experiments

In this section we perform several experiments that are meant to illustrate how our model economy behaves and compares with the data. We first show the response of the economy to a capital quality shock with and without runs to illustrate how the model generates a financial panic. We then compare how runs versus occasionally binding constraints can generate nonlinear dynamics. Finally, we turn to an experiment that shows how the model can replicate salient features of the recent financial crisis.

4.2.1 Response to a capital quality shock: no bank run case

We suppose the economy is initially in a risk-adjusted steady state. Figure 1 shows the response of the economy to a negative one standard deviation (.75%) shock to the quality of capital.[18] The solid line is our baseline model and the dotted line is the case where financial frictions are shut off. For both cases the shock reduces the expected return to capital, reducing investment and in turn aggregate demand. In addition, for the baseline economy with financial frictions, the weakening of bank balance sheets amplifies the contraction in demand through the financial accelerator or credit cycle mechanism of Bernanke Gertler and Gilchrist (1999) and Kiyotaki and Moore(1997). Poor asset returns following the shock cause bank net worth to decrease by about 15%. As bank net worth declines, incentive constraints tighten and banks decrease their demand for assets causing the price of capital to drop. The

[18] In all of the experiments we trace the response of the economy to the shocks considered assuming that after these shocks capital quality is exactly equal to its conditional expectations, i.e. setting future ε_t to 0.

drop in asset prices feeds back into lower bank net worth, an effect that is magnified by the extent of bank leverage. As financial constraints tighten and asset prices decline, excess returns rise by 75 basis points which allows banks to increase their leverage by about 10%. Overall, a 0.75 percent decline in the quality of capital results in a drop in investment by 5 percent and a drop in output by slightly more than 1 percent. The drop in investment is roughly double the amount in the case absent financial frictions, while the drop in output is about thirty percent greater.

In the experiment of Figure 1, the economy is always ex post in a "safe zone", where a bank run equilibrium does not exist. Under our parametrization, a bank run cannot happen in the risk-adjusted steady state: bank leverage is too low. The dashed line in the first panel of Figure 1 shows the size of the shock in the subsequent period needed to push the economy into the run region. In our example, a two standard deviation shock is needed to open up the possibility of runs starting from the risk adjusted steady state, which is double the size of the shock considered in Figure 1.

Even though in this case the economy is always in a safe region ex post, it is possible ex ante that a run equilibrium could occur in the subsequent period. In particular, the increase in leverage following the shock raises the probability that a sufficiently bad shock in the subsequent period pushes the economy into the run region. As the top middle panel of Figure 1 shows, the overall probability of a run increases following the shock.

4.2.2 Bank runs

In the previous experiment the economy was well within a safe zone. A one standard deviation shock did not and could not produce a financial panic. We now consider a case where the economy starts in the safe zone but is gradually pushed to the edge of the crisis zone, where runs are possible. We then show how a shock of the same magnitude as in Figure 1 can induce a panic with damaging effects on the real economy.

To implement this experiment, we assume that the economy is hit by a sequence of three equally sized negative shocks that push the economy to the run threshold. That is, we find a shock ϵ^* that satisfies:

$$\xi_3^R = \epsilon^* \left(1 + \rho_\xi + \rho_\xi^2\right)$$

where ξ_3^R is the threshold level for the capital quality below which a run is possible at time 3, given that the economy is in steady state at time 0 and is hit by two equally sized shocks at time 1 and 2, i.e. $\epsilon_1 = \epsilon_2 = \epsilon^*$. The first two shocks push the economy to the edge of the crisis zone. The third pushes it just in.

The solid line in Figure 2 shows the response of the economy starting from period two onwards under the assumption that the economy experiences a run with arrival of a sunspot in period three. For comparison, the dashed line shows the response of the economy to the same exact capital quality shocks but assuming that no sunspot is observed and so no run happens.

As shown in panel 1 the size of the threshold innovation of capital quality shock turns out to be roughly equal to one standard deviation, i.e. .83%., which is the size of the shock in Figure 1. After the first two innovations, the capital quality is 1.4% below average and the run probability is about 2% quarterly. The last innovation pushes the economy into the run region. When the run happens, bank net worth is wiped out which forces banks to liquidate assets. In turn, households absorb the entire capital stock. Households however are only willing to increase their portfolio holdings of capital at a discount, which leads excess returns to spike and investment to collapse. When the run occurs, investment drops an additional 25% resulting in an overall drop of 35%. Comparing with the case of no run clarifies that almost none of this additional drop is due to the capital quality shock itself: The additional drop in investment absent a run is only 2.5%. The collapse in investment demand causes inflation to decrease and induces monetary policy to ease by reducing the policy rate to slightly below zero . However, reducing the nominal interest rate to roughly zero is not sufficient to insulate output which drops 7%.

As new bankers enter the economy, bank net worth is slowly rebuilt and the economy returns to the steady state. This recovery is slowed down by a persistent increase in the run probability following the banking panic. The increase in the run probability reduces the amount of leverage that banks are willing to take on.

To get a sense of the role that nominal rigidities are playing, Figure 3 describes the effect of bank runs in the economy with flexible prices. For comparison, with the analogous experiment in our baseline (in Figure 2) we hit the flex price economy with the same sequence of shocks that would take the baseline economy to the run threshold.[19] There are two main takeaways

[19]However, since in the flex price economy there is much less amplification, the ex-post run that we consider is actually not an equilbrium. As the first panel in the figure

from Figure 3. First, the output drop in the flexible price case is only about half that in our benchmark sticky price case. The New Keynesian features thus magnify the effects of the banking crisis. The reason is that the banking crisis generates a steep decline in the natural rate of interest by inducing a collapse in investment demand. As a result, in the flexible price case the real interest rate drops roughly eight hundred basis points below zero leading to a temporary expansion in consumption demand and hence dampening the output contraction. Such a dramatic drop in real rates is clearly not be feasible with nominal rigidities and a zero lower bound. Second, even in the flexible price case, a bank run will amplify the contraction in output by inducing a large drop in investment demand. In our example, relative to the no run case. the run increases the drop in output from about one percent to three and a half percent.

4.2.3 Nonlinearities: occasionally binding constraints vs runs

We now turn to nonlinearities within our baseline model. We will start by considering the effects of occasionally binding constraints. Figure 4 shows the behavior of the economy when it transits from slack to binding financing (incentive) constraints. Starting from the risk adjusted steady state we consider how the economy responds to variations in capital quality. If the shock to capital quality is positive the constraint is slack, while it becomes binding with negative capital quality shocks. Overall, nonlinearities are present, though they do not turn out to be as large as in the case of bank runs. A negative capital quality shock affects investment, asset prices and credit spreads only a little more, in absolute value, than does a similar magnitude increase. The asymmetries arising in our framework are somewhat dampened for two reasons: First, in many frameworks the maximum feasible leverage multiple is fixed (e.g. Mendoza, 2010). However, in our model, as the economy moves into the constrained region the maximum feasible leverage multiple increases (see section 2.2.2). This relaxing of the leverage constraint reduces the decline in real activity and asset prices and the rise in credit spreads. Second, it is often assumed that the real interest rate is fixed. In our model, however, the real rate declines as the economy weakens, which also works to dampen the decline in the constrained region.

shows, even after the first two shocks the shock that is needed to push the economy to the threshold is still very large in the flex price economy, i.e. around -4%.

Next we consider bank runs. Figure 5 shows the response of the economy to a capital quality shock starting from the same initial state considered in Figure 2. The dashed line depicts the response in the case in which no sunspot occurs (so that a bank run cannot happen) and the solid one shows the case in which a sunspot is realized (so that a run will occur if a run equilibrium exists). As long as capital quality shocks are above the run threshold the responses are identical in the two cases since in this region a run is not possible. When the shock lies below the run threshold, however, a run equilibrium is possible. In this region, when agents observe a sunspot they run on financial institutions pushing the economy to an equilibrium in which banks are forced to liquidate assets at fire sale prices. The highly nonlinear behavior we described in the introduction then emerges: excess returns spike and investment and asset prices collapse.

4.2.4 Crisis experiment: model versus data

Figure 6 illustrates how the model can replicate some salient features of the recent financial crisis. We hit the economy with a series of capital quality shocks over the period 2007Q4 until 2008Q3. The starting point is the beginning of the recession, which roughly coincides with the time credit markets first came under stress following Bear Stearns' losses on its MBS portfolios. We pick the size of the capital quality shocks to match the observed decline in investment during this period, in panel 1. We then assume that a run happens in 2008Q4, the quarter in which Lehman failed and the shadow banking system collapsed. The solid line shows the observed response of some key macroeconomic variables.[20] The dashed line shows the response of the economy when a run occurs in 2008Q4 and the dotted line shows the response under the assumption that a run does not happen.

As indicated in the figure, the sequence of negative surprises in the quality of capital needed to match the observed contraction in investment leads to a gradual decline in banks net worth that matches closely the observed decline in financial sector equity as measured by the XLF index, which is an index of S&P 500 financial stocks, in panel 2. Given that banks net worth is already depleted by poor asset returns, a very modest innovation in 2008Q4 pushes the economy into the run region. When the run occurs, the model economy

[20]For output, investment and consumption we show deviation from a trend computed by using CBO estimates of potential output and similarly for hours worked we let the CBO estimate of potentail labor represent the trend.

generates a sudden spike in excess returns and a drop in investment, output, consumption and employment of similar magnitudes as those observed during the crisis. The dotted line shows how, absent a run, the same shocks would generate a much less severe downturn.

The model economy also predicts a rather slow recovery following the financial crisis, although faster than what we observed in the data. It is important however to note that in the experiment we are abstracting from any disturbances after 2008Q4. This implies a rather swift recovery of financial equity and excess returns to their long run value. On the other hand, the observed recovery of net worth and credit spreads was much slower with both variables still far from their pre-crisis values as of today. Various factors that are not captured in our model economy, such as a drastic change in the regulatory framework of financial institutions, increased uncertainty following the crisis and slow adjustment of household balance sheets, have likely contributed to the very slow recovery of these financial variables. Incorporating these factors could help the model account for the very slow recovery of investment and employment. However we leave this extension for future research.

5 Conclusion

We have developed a macroeconomic model with a banking sector where costly financial panics can arise. A panic or run in our model is a self-fulfilling failure of creditors to roll over their short-term credits to banks. When the economy is close to the steady state a self-fulfilling rollover crises cannot happen because banks have sufficiently strong balance sheets. In this situation, "normal size" business cycle shocks do not lead to financial crises. However, in a recession, banks may have sufficiently weak balance sheet so as to open up the possibility of a run. Depending on the circumstances either a small shock or no further shock can generate a run that has devastating consequences for the real economy. We show that our model generates the highly nonlinear contraction in economic activity associated with financial crises. It also captures how crises may occur even in the absence of large exogenous shocks to the economy. We then illustrate that the model is broadly consistent with the recent financial crisis.

One issue we save for further work is the role of macroprudential policy. As with other models of macroprudential policy, externalities are present that lead banks to take more risk than is socially efficient. Much of the literature is based on the pecuniary externality analyzed by Lorenzoni (2008), where individual banks do not properly internalize the exposure of the system to asset price fluctuations that generate inefficient volatility, but not runs. A distinctive feature of our model is that the key externality works through the effect of leverage on the bank run probability: Because the run probability depends on the leverage of the banking system as a whole, individual banks do not fully take into account the impact of their own leverage decisions on the exposure of the entire system. In this environment, the key concern of the macroprudential policy becomes reducing the possibility of a financial collapse in the most efficient way. Our model will permit us to explore the optimal design of policies qualitatively and quantitatively.

References

[1] Adrian, T., and H. Shin. 2010. The Changing Nature of Financial Intermediation and the Financial Crisis of 2007-2009. *Annual Review of Economics*: 603-618.

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

[3] Benhabib, J., and P. Wang. 2013. Financial Constraints, Endogenous Markups, and Self-Fulfilling Equilibria. *Journal of Monetary Economics* 60: 789-805

[4] Bernanke, B., 2010. Causes of the Recent Financial and Economic Crisis. Statement before the Financial Crisis Inquiry Commission, Washington, September 2.

[5] Bernanke, B., M. Gertler and S. Gilchrist. 1999. Financial Accelerator in a Quantitative Business Cycle Framework. In Taylor, J.B., and Woodford, M. (Eds.), *Handbook of Macroeconomics.* Vol. 1: 1341-1393. Elsevier, Amsterdam, Netherlands.

[6] Bocola L. and G. Lorenzoni, 2017, "A Model of Financial Crises in Open Economies," mimeo.

[7] Bordo,M. D., B. Eichengreen, D. Klingebiel and M. Soledad Martinez-Peria. 2001. Is the Crisis Problem Growing More Severe?. *Economic Policy: A European Forum* 32: 51-75.

[8] Brunnermeier, M. K., T. Eisenbach and Y. Sannikov. 2013. Macroeconomics with Financial Frictions: A Survey." In D. Acemoglu, M. Arellano and E. Dekel (Eds.) *Advances in Economics and Econometrics: Tenth World Congress of the Econometric Society.* Vol. 2: 3-94. Cambridge: Cambridge University Press.

[9] Brunnermeier, M. K., and Y. Sannikov. 2014. A Macroeconomic Model with a Financial Sector. *American Economic Review* 104, 379-421.

[10] Calvo, G. 1988. "Servicing the Public Debt: The Role of Expectations, *American Economic Review* 78, 647-661.

[11] Cole, H. and T. Kehoe, 2000, Self-fulfilling Debt Crises, *Review of Economic Studies* 67, 91-161.

[12] Cooper, R. and T. Ross, 1998, Bank Runs: Liquidity Costs and Investment Distortions, *Journal of Monetary Economics* 41, 27-38.

[13] Del Negro, M., M. Giannoni and F. Shorfheide 2015. Inflation in the Great Recession and New Keynesian Models. *American Economic Journal*, 7(1): 168-196

[14] Diamond, D., and P. Dybvig. 1983. Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy* 91, 401-419.

[15] Farhi, E. and M. Maggiori, 2017, "A Model of the International Monetary System", *Quarterly Journal of Economics,* forthcoming.

[16] Gertler, M., and N. Kiyotaki. 2011. Financial Intermediation and Credit Policy in Business Cycle Analysis. In Friedman, B., and Woodford, M. (Eds.), *Handbook of Monetary Economics.* Vol. 3A: 547-599. Elsevier, Amsterdam, Netherlands.

[17] Gertler, M., N. Kiyotaki. 2015. Banking, Liquidity and Bank Runs in an Infinite Horizon Economy, *American Economic Review,* 105: 2011-2043.

[18] Gertler, M., N. Kiyotaki and A. Prestipino. 2016. Wholesale Banking and Bank Runs in Macroeconomic Modeling. In Taylor, J.B., and H. Uhlig. (Eds.), *Handbook of Macroeconomics.* Vol. 2B: 1345-1425. Elsevier, Amsterdam, Netherlands.

[19] Gilchrist, S., and E. Zakrajsek. 2012. Credit Spread and Business Cycle Fluctuations, *American Economic Review*, 102: 1692-1720.

[20] Gorton, G., 2010. *Slapped by the Invisible Hand: The Panic of 2007.* Oxford: Oxford University Press.

[21] He, Z. and A. Krishnamurthy, 2017. A Macroeconomic Framework for Quantifying Systemic Risk. Working Paper, University of Chicago and Stanford University.

[22] He, Z., I.G. Khang, and A. Krishnamurthy, 2010. Balance Sheet Adjustment during the 2008 Crisis. *IMF Economic Review* 1: 118-156.

[23] Kiyotaki, N., and J. Moore. 1997. Credit Cycles. *Journal of Political Economy* 105, 211-248.

[24] Kurlat, P. 2016. Asset markets With Heterogeneous Information. *Econometrica*. 84(1):33-85.

[25] Lorenzoni, G., 2008. Inefficient Credit Booms. *Review of Economic Studies* 75, 809-833.

[26] Lorenzoni, G., and K. Wallentin. 2007. Financial Frictions, Investment and Tobin's q. Mimeo, Northwestern University.

[27] Martin, A., D. Skeie and E.L. Von Thadden. Repo Runs. *Review of Financial Studies*, 27: 957-989.

[28] Mendoza, E. 2010. "Sudden Stops, Financial Crises, and Leverage. *American Economic Review* 100, 1941-1966.

[29] Perri, F. and V. Quadrini. forthcoming. "International Recessions," *American Economic Review.*

[30] Phillipon, T. 2015. Has the U.S. Finance Industry Become Less Efficient? *American Economic Review.*

[31] Reinhart, C. and K. Rogoff 2009. This Time Is Different: Eight Centuries of Financial Folly. Princeton,NJ: Princeton University Press.

[32] Robatto, R. 2014, Financial Crises and Systematic Bank Runs in a Dynamic Model of Banking, Working Paper, University of Wisconsin-Madison

[33] Schularick, M. and A. Taylor. 2012. Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008. *American Economic Review*102(2): 1029-1061

[34] Uhlig, H., 2010, A Model of a Systemic Bank Run, *Journal of Monetary Economics* 57, 78-96.

Table 1

Parameter	Description	Value	Target
	Standard Parameters		
β	Impatience	.99	Risk Free Rate
γ_h	Risk Aversion	2	Literature
φ	Frish Elasticity	2	Literature
ϵ	Elasticity of subst across varieties	11	Markup 10%
α	Capital Share	.33	Capital Share
δ	Depreciation	.025	$\frac{I}{K} = .025$
η	Elasticity of q to i	.25	Literature
a	Investment Technology Parameter	.53	$Q = 1$
b	Investment Technology Parameter	-.83%	$\frac{I}{K} = .025$
G	Government Expenditure	.45	$\frac{G}{Y} = .2$
ρ^{jr}	Price adj costs	1000	Slope of Phillips curve .01
κ_π	Policy Response to Inflation	1.5	Literature
κ_y	Policy Response to Output	.5	Literature
	Financial Intermediation Parameters		
σ	Banker Survival rate	.93	Leverage $\frac{QS^b}{N} = 10$
ζ	New Bankers Endowments as a share of Capital	.1%	% Δ I in crisis $\approx$ 35%
θ	Share of assets divertible	.23	Spread Increase in Crisis = 1.5%
γ	Threshold for HH Intermediation Costs	.432	$\frac{S^b}{S} = .5$
χ	HH Intermediation Costs	.065	$ER^b - R = 2\%$ Annual
$\varkappa$	Sunspot Probability	.15	Run Probability 4% Annual
$\sigma(\epsilon^\xi)$	std of innovation to capital quality	.75%	std Output
ρ^ξ	serial correlation of capital quality	.7	std Investment

Table 2: Standard Deviations Data vs. Model

	Data 1983–2007q3	Model No Runs Happen
Y	1.9	2.4
C+I	2.7	3.0
I	7.2	6.9
C	1.3	3.1
L	3.1	3.4

All values in percentages.

NOTE: For output, investment, consumption, and government spending we compute real per capita terms by dividing the nominal variables by the population and adjusting by the GDP deflator. For labor we compute per capita hours worked by dividing total labor hours by the population. We then show the standard deviations of the logged variables in deviations from a linear trend starting in 1983q1 and ending in 2007q3.

SOURCE: Output, investment (gross private domestic investment plus durable good consumption), consumption (personal consumption expenditure less durable good consumption), government spending, and the GDP deflator are from the Bureau of Economic Analysis. Total labor hours (aggregate hours, nonfarm payrolls) and population (civilian noninstitutional, 16 years and over) are from the Bureau of Labor Statistics.

Fig. 1. Response to a Capital Quality Shock (1 std): No Run Case

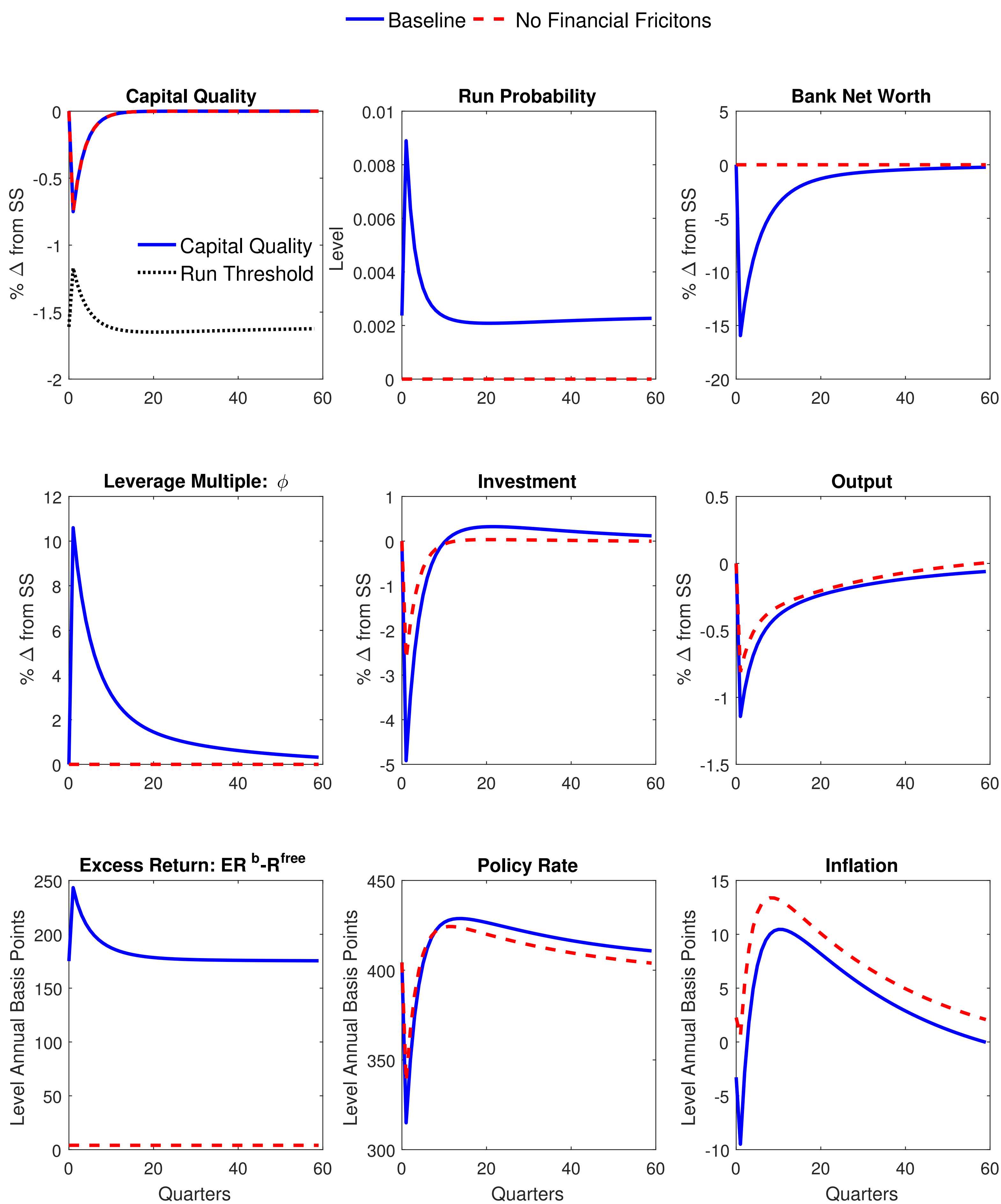


Fig. 2. Response to a Sequence of Shocks: Run VS No Run

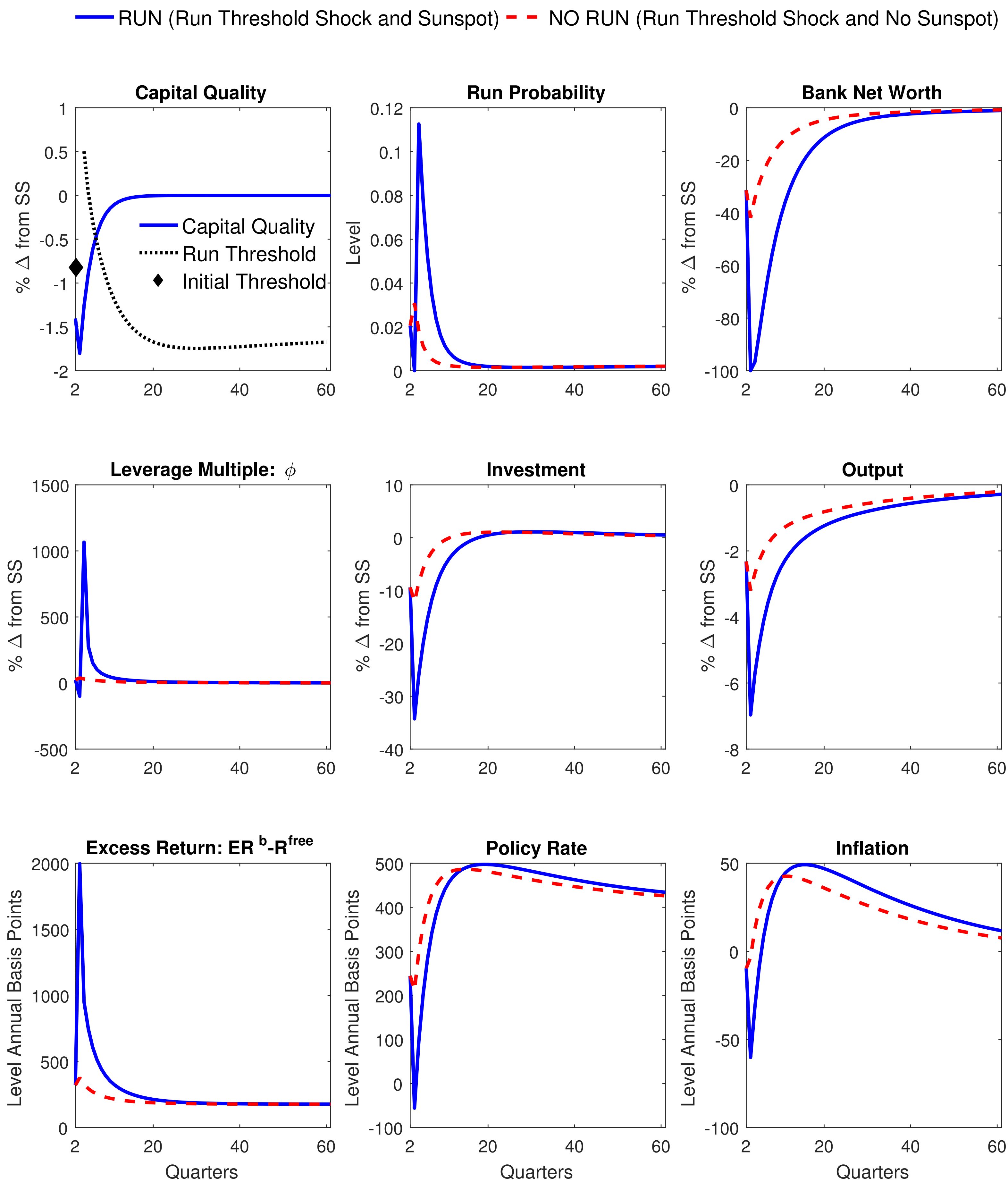


Fig. 3. Response to the Same Sequence of Shocks in Flex Price Economy: Run VS No Run

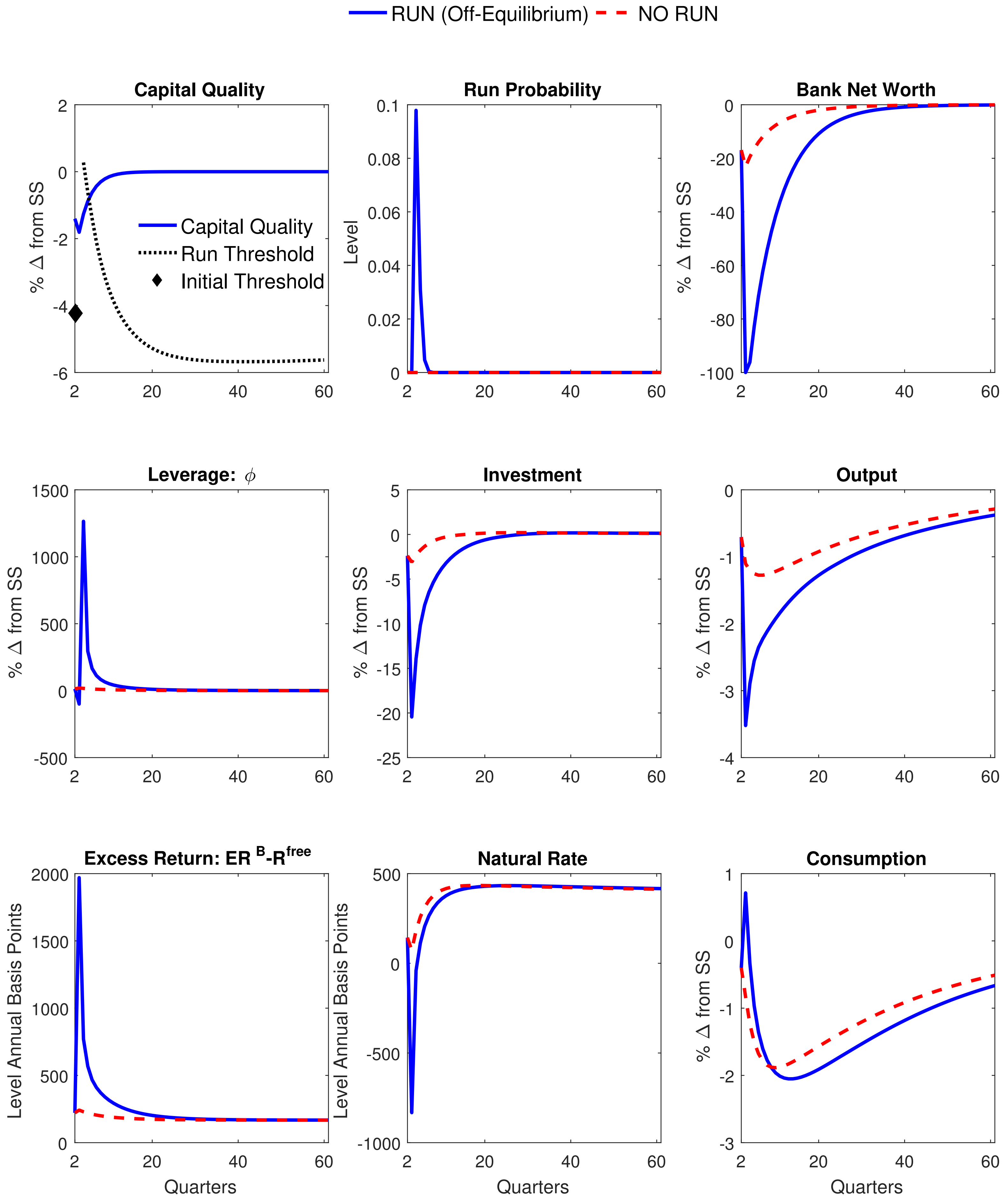


Fig. 4. Non-Linearities due to Occasionally Binding Constraints

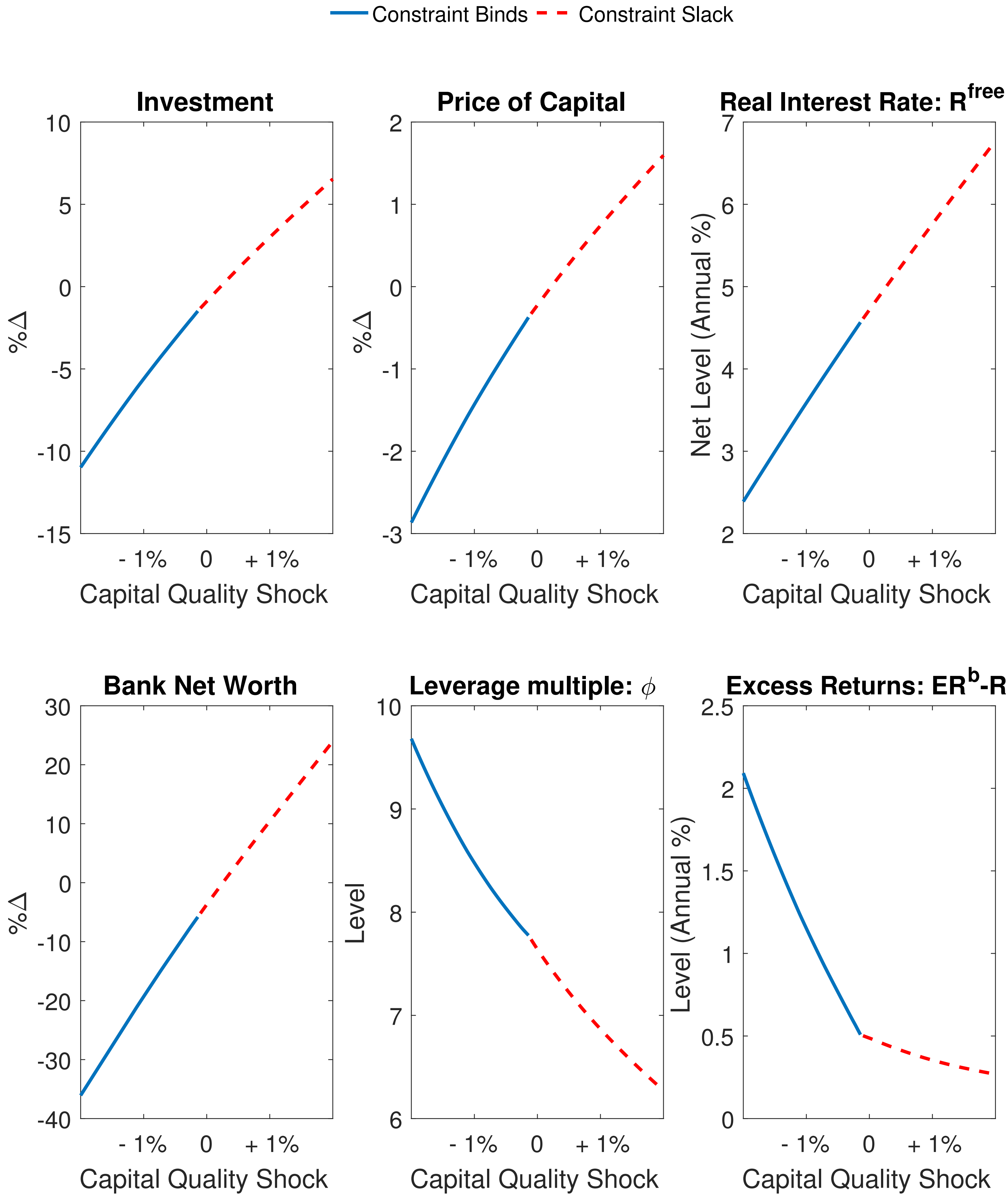


Fig. 5. Non-linearities from Runs

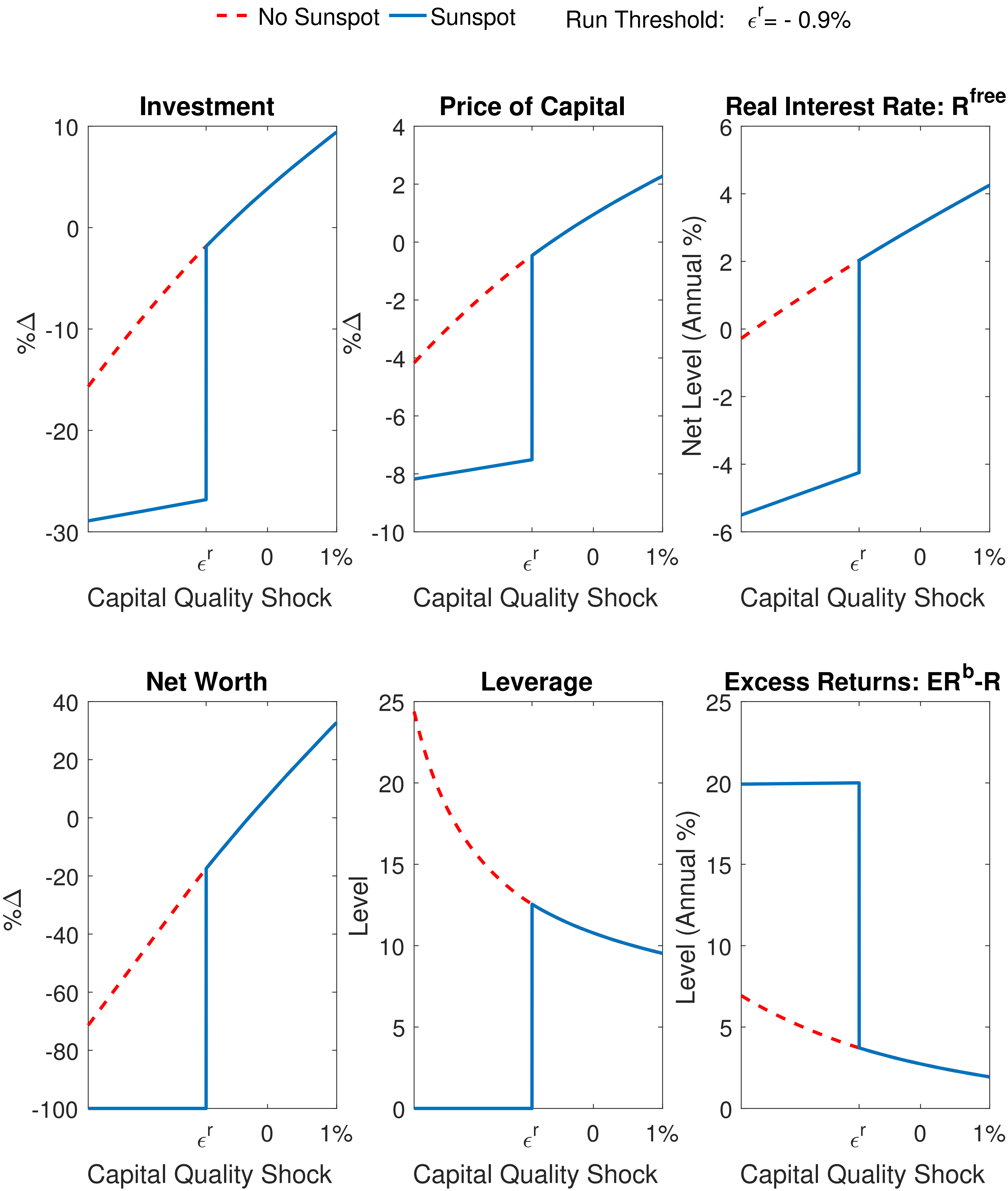
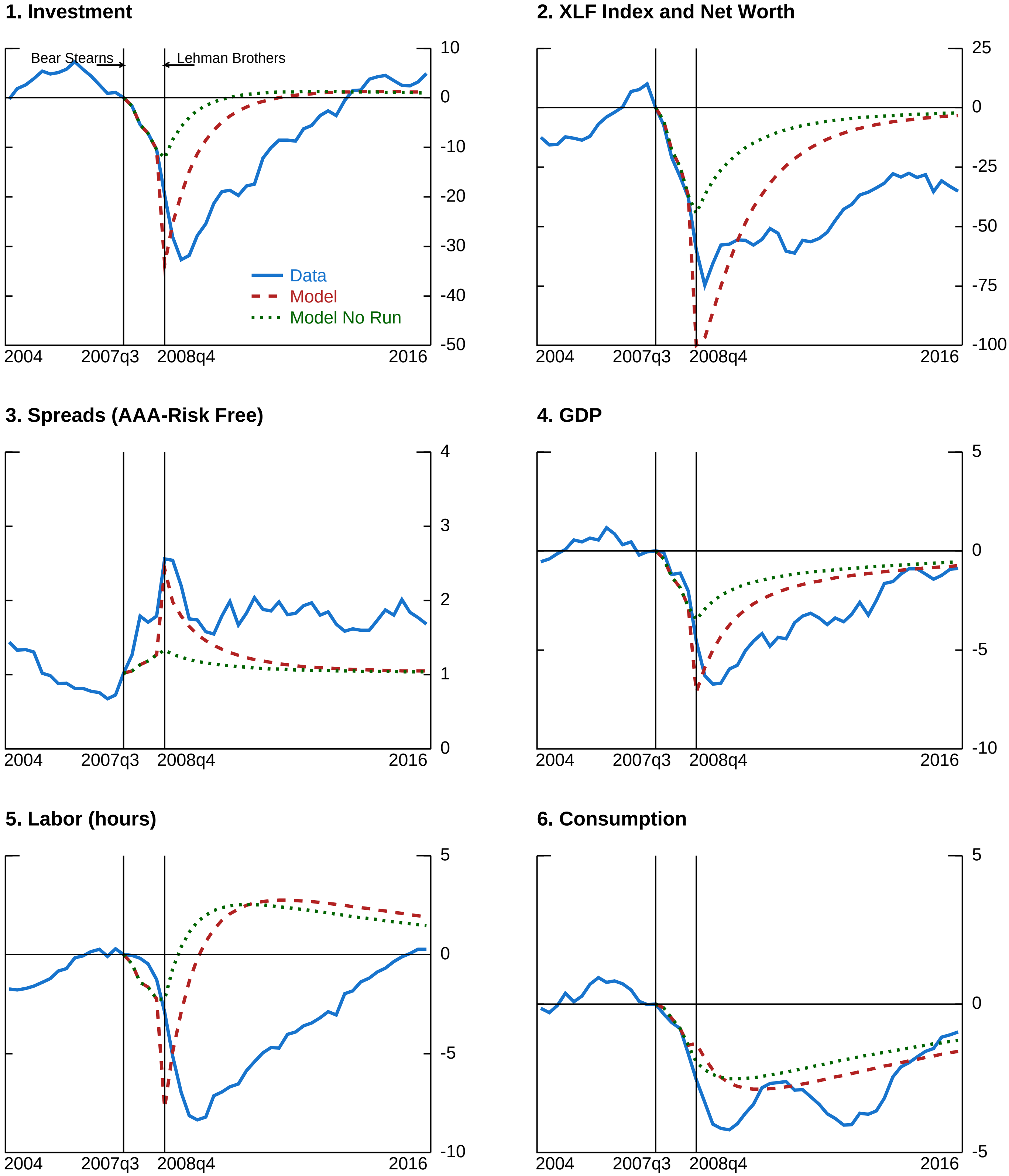


Fig. 6. Financial Crisis: Model vs. Data

Shocks* :	-0.3 %	-0.6 %	-0.5 %	-0.8 %	-0.7 %
	2007q4	2008q1	2008q2	2008q3	2008q4



NOTE: The data for GDP, Investment, and Consumption are computed as logged deviations from trend where the trend is the CBO potential GDP. Labor data is computed as logged deviations from trend where the trend is the CBO potential hours worked. The XLF Index data is computed as the percent deviation from its 2007q3 level.

6 Appendix

This Appendix describes the details of the equilibrium.

The aggregate state of the economy is summarized by the vector $\tilde{\mathcal{M}}_t = (S_{t-1}, S^b_{t-1}, \overline{R}_t D_{t-1}, \xi_t, \iota_t)$, with: S_{t-1} = capital stock at the end of $t-1$; S^b_{t-1} = bank capital holdings in $t-1$; $\overline{R}_t D_{t-1}$ = bank deposit obligation at the beginning of t; ξ_t = capital quality shock realized in t; and ι_t =sunspot realization at time t.

6.1 Producers

As described in the text, the capital stock for production in t is given by

$$K_t = \xi_t S_{t-1}, \tag{54}$$

The capital quality shock is serially correlated as follows

$$\xi_{t+1} \sim F\left(\xi_{t+1} \mid \xi_t\right) = F_t\left(\xi_{t+1}\right)$$

with a continuous density:

$$F_t'\left(\xi_{t+1}\right) = f_t\left(\xi_{t+1}\right), \text{ for } \xi_{t+1} \in (0, \infty).$$

Capital at the end of period is

$$S_t = \Gamma\left(\frac{I_t}{K_t}\right) K_t + (1-\delta) K_t. \tag{55}$$

As we described in the text, capital goods producer's first order condition for investment is

$$Q_t \Gamma'\left(\frac{I_t}{K_t}\right) = 1. \tag{56}$$

A final goods firms chooses intermediate goods $\{Y_t(f)\}$ to minimize the cost

$$\int_0^1 P_t(f) Y_t(f) df$$

subject to the production function:

$$Y_t = \left[\int_0^1 Y_t(f)^{\frac{\varepsilon-1}{\varepsilon}} df\right]^{\frac{\varepsilon}{\varepsilon-1}}. \tag{57}$$

Cost minimization then yields a demand function for each intermediate good f :

$$Y_t(f) = \left[\frac{P_t(f)}{P_t}\right]^{-\varepsilon} Y_t, \tag{58}$$

where P_t is the price index, given by

$$P_t = \left[\int_0^1 P_t(f)^{1-\varepsilon} df\right]^{\frac{1}{1-\varepsilon}}.$$

Conversely, an intermediate goods producer f chooses input to minimize the production cost

$$w_t L_t(f) + Z_t K_t(f)$$

subject to

$$A_t[K_t(f)]^{\alpha}[L_t(f)]^{1-\alpha} = Y_t(f).$$

The first order conditions yield

$$\frac{K_t(f)}{L_t(f)} = \frac{\alpha}{1-\alpha}\frac{w_t}{Z_t} = \frac{K_t}{L_t}, \tag{59}$$

and the following relation for marginal cost:

$$MC_t = \frac{1}{A_t}\left(\frac{Z_t}{\alpha}\right)^{\alpha}\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha}. \tag{60}$$

Each period, the intermediate goods producer chooses $P_t(f)$ and $Y_t(f)$ to maximize the expected discounted value of profits:

$$E_t\left\{\sum_{\tau=t}^{\infty}\widetilde{\Lambda}_{t,\tau}\left[\left(\frac{P_\tau(f)}{P_\tau} - MC_\tau\right)Y_\tau(f) - \frac{\rho^r}{2}Y_\tau\left(\frac{P_\tau(f)}{P_{\tau-1}(f)} - 1\right)^2\right]\right\},$$

subject to the demand curve (58), where $\widetilde{\Lambda}_{t,\tau} = \beta^{\tau-t}(C_\tau/C_t)^{-\gamma_h}$ is the discount factor of the representative household. Taking the firm's first order condition for price adjustment and imposing symmetry implies the following forward looking Phillip's curve:

$$(\pi_t - 1)\pi_t = \frac{\varepsilon}{\rho^r}\left(MC_t - \frac{\varepsilon-1}{\varepsilon}\right) + E_t\left[\widetilde{\Lambda}_{t,t+1}\frac{Y_{t+1}}{Y_t}(\pi_{t+1}-1)\pi_{t+1}\right]. \tag{61}$$

where $\pi_t = \frac{P_t}{P_{t-1}}$ is the realized gross inflation rate at date t. The cost minimization conditions with symmetry also imply that aggregate production is simply

$$Y_t = A_t {K_t}^{\alpha} {L_t}^{1-\alpha}. \tag{62}$$

6.2 Households

We modify the household's maximization problem in the text by allowing for a riskless nominal bond which will be in zero supply. We do so to be able the pin down the riskless nominal rate R_t^n. Let B_t be real value of this riskless bond. The household then chooses C_t, L_t, B_t, D_t and S_t^h to maximize expected discounted utility U_t:

$$U_t = E_t \left\{ \sum_{\tau=t}^{\infty} \beta^{\tau-t} \left[\frac{(C_\tau)^{1-\gamma_h}}{1-\gamma_h} - \frac{(L_\tau)^{1+\varphi}}{1+\varphi} - \varsigma(S_\tau^h, S_\tau) \right] \right\},$$

subject to the budget constraint

$$C_t + D_t + Q_t S_t^h + B_t = w_t L_t - T_t + \Pi_t + R_t D_{t-1} + \frac{R_{t-1}^n}{\pi_t} B_{t-1} + \xi_t [Z_t + (1-\delta) Q_t] S_{t-1}^h.$$

As explained in the text, the rate of return on deposits is given by

$$\begin{aligned} R_t &= Max \left\{ \overline{R}_t, \frac{\xi_t [Z_t + (1-\delta) Q_t] S_{t-1}^b}{D_{t-1}} \right\} \\ &= Max \left\{ \overline{R}_t, \frac{\xi_t [Z_t + (1-\delta) Q_t]}{Q_{t-1}} \frac{Q_{t-1} S_{t-1}^b}{Q_{t-1} S_{t-1}^b - N_{t-1}} \right\} \\ &= Max \left(\overline{R}_t, R_t^b \frac{\phi_{t-1}}{\phi_{t-1} - 1} \right), \end{aligned}$$

where $R_t^b = \frac{\xi_t [Z_t + (1-\delta) Q_t]}{Q_{t-1}}$ and where $\phi_t = Q_t S_t^b / N_t$ is the bank leverage multiple.

We obtain the first order conditions for labor, riskless bonds, deposits and direct capital holding, as follows:

$$w_t = (C_t)^{\gamma_h} (L_t)^{\varphi} \tag{63}$$

$$E_t \left(\Lambda_{t+1} \frac{R_t^n}{\pi_{t+1}} \right) = 1 \tag{64}$$

$$E_t \left[\Lambda_{t+1} Max \left(\overline{R}_{t+1}, R_{t+1}^b \frac{\phi_t}{\phi_t - 1} \right) \right] = 1 \tag{65}$$

$$E_t \left\{ \Lambda_{t+1} \xi_{t+1} \frac{Z_{t+1} + (1-\delta) Q_{t+1}}{Q_t + \frac{\partial}{\partial S_t^h} \varsigma(S_t^h, S_t) \cdot C_t{}^{\gamma_h}} \right\} = 1, \tag{66}$$

where

$$\Lambda_{t+1} = \widetilde{\Lambda}_{t,t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma_h}, \text{ and}$$

$$\frac{\partial}{\partial S_t^h} \varsigma(S_t^h, S_t) = Max \left[\chi \left(\frac{S_t^h}{S_t} - \gamma \right),\ 0 \right].$$

6.3 Bankers

For ease of exposition, the description of the banker's problem in the text does not specify how the individual choice of bank's leverage affects its own probability of default. This was possible because, as argued in footnote 7, the marginal effect of leverage on the objective of the firm, V_t, through the change in p_t is zero. Therefore the first order conditions for the bank's problem, equations (25) and (26), can be derived irrespectively of how the individual choice of bank's leverage affects its own probability of default.

We now formalize the argument in footnote 7 and describe how the default thresholds for individual banks vary with individual bank leverage. As will become clear in section 6.5 below, this analysis is key in order to study global optimality of the leverage choice selected by using the first order conditions in the text, equations (25) and (26).

As in the text, ι is a sunspot which takes on values of either unity or zero. We can then express the rate of return on bank capital R_{t+1}^b

$$R_{t+1}^b = \xi_{t+1} \frac{Z_{t+1} + (1-\delta) Q_{t+1}}{Q_t} = R_{t+1}^b(\xi_{t+1}, \iota_{t+1}),$$

The individual bank defaults at date t+1 if and only if

$$1 > \frac{\xi_{t+1}[Z_{t+1} + (1-\delta)Q_{t+1}] s_t^b}{\overline{R}_{t+1} d_t} = \frac{R_{t+1}^b(\xi_{t+1}, \iota_{t+1})}{\overline{R}_{t+1}} \frac{Q_t s_t^b}{Q_t s_t^b - n_t},$$

or

$$R_{t+1}^b(\xi_{t+1}, \iota_{t+1}) < \overline{R}_{t+1} \frac{\phi_t}{\phi_t - 1}.$$

Let $\Xi_{t+1}^D(\phi)$ be the set of capital quality shocks and sunspot realizations which make the individual bank with a leverage multiple of ϕ default and conversely let $\Xi_{t+1}^N(\phi)$ be the set that leads to non-default at date t+1:

$$\Xi_{t+1}^D(\phi) = \left\{ (\xi_{t+1}, \iota_{t+1}) \mid R_{t+1}^b(\xi_{t+1}, , \iota_{t+1}) < \frac{\phi - 1}{\phi} \overline{R}_{t+1}(\phi) \right\},$$

$$\Xi^N_{t+1}(\phi) = \left\{ (\xi_{t+1}, \iota_{t+1}) \mid R^b_{t+1}(\xi_{t+1}, \iota_{t+1}) \geq \frac{\phi - 1}{\phi} \overline{R}_{t+1}(\phi) \right\}.$$

where $\overline{R}_{t+1}(\phi)$ is the promised deposit interest rate when the individual bank chooses ϕ which satisfies the condition for the household to hold deposits:

$$1 = \overline{R}_{t+1}(\phi) \int_{\Xi^N_{t+1}(\phi)} \Lambda_{t+1} d\widetilde{F}_t + \frac{\phi}{\phi - 1} \int_{\Xi^D_{t+1}(\phi)} \Lambda_{t+1} R^b_{t+1}(\xi_{t+1}, \iota_{t+1}) d\widetilde{F}_t. \quad (67)$$

Here $\widetilde{F}_t(\xi_{t+1},, \iota_{t+1})$ denotes the distribution function of (ξ_{t+1}, ι_{t+1}) conditional on date t information.

Assume that the aggregate leverage multiple is given by $\overline{\phi}_t$. When the individual banker chooses the leverage multiple ϕ, which can be different from $\overline{\phi}_t$, the individual bank defaults at date t+1 if and only if

$$\begin{gathered} \xi_{t+1} < \xi^I_{t+1}(\phi) \text{ and } \iota_{t+1} = 0 \\ where \\ \xi^I_{t+1}(\phi) R^b_{t+1}(\xi^I_{t+1}(\phi), 0) = \tfrac{\phi-1}{\phi} \overline{R}_{t+1}(\phi) \end{gathered} \quad (68)$$

or

$$\begin{gathered} \xi_{t+1} < \xi^R_{t+1}(\phi) \text{ and } \iota_{t+1} = 1 \\ where \\ \xi^R_{t+1}(\phi) = \sup \left\{ \xi_{t+1} \text{ s.t. } \xi_{t+1} R^b_{t+1}(\xi_{t+1}, 1) < \tfrac{\phi-1}{\phi} \overline{R}_{t+1}(\phi) \right\}. \end{gathered} \quad (69)$$

Thus the set of capital quality shocks and sunspots which make the individual bank default $\Xi^D_{t+1}(\phi)$ is

$$\Xi^D_{t+1}(\phi) = \left\{ (\xi_{t+1}, \iota_{t+1}) \;\middle|\; \begin{array}{c} \xi_{t+1} < \xi^I_{t+1}(\phi) \text{ and } \iota_{t+1} = 0 \\ \text{or} \\ \xi_{t+1} < \xi^R_{t+1}(\phi) \text{ and } \iota_{t+1} = 1 \end{array} \right\}. \quad (70)$$

$$\Xi^N_{t+1}(\phi) = \left\{ (\xi_{t+1}, \iota_{t+1}) \;\middle|\; \begin{array}{c} \xi_{t+1} \geq \xi^I_{t+1}(\phi) \text{ and } \iota_{t+1} = 0 \\ \text{or} \\ \xi_{t+1} \geq \xi^R_{t+1}(\phi) \text{ and } \iota_{t+1} = 1 \end{array} \right\}. \quad (71)$$

The behavior of $\xi^I_{t+1}(\phi)$ is straightforward and can be easily characterized from (68) under the natural assumption that R^b_{t+1} is increasing in the quality of capital at $t + 1$. This gives:

$$\begin{gathered} \tfrac{d\xi^I_{t+1}(\phi)}{d\phi} > 0, \text{ for } \phi \in (1, \infty) \\ \lim_{\phi \downarrow 1} \xi^I_{t+1}(\phi) = 0. \end{gathered} \quad (72)$$

The behavior of $\xi^R_{t+1}(\phi)$ is more complicated because, when a sunspot is observed, the function $R^b_{t+1}\left(\xi_{t+1},1\right)$ that determines returns on bank's assets as a function of the capital quality is discontinuous around the aggregate run threshold $\xi^R_{t+1} = \xi^R_{t+1}(\bar{\phi}_t)$: at the threshold ξ^R_{t+1} asset prices jump from liquidation prices up to their normal value (See Figure 5):

$$\lim_{\xi_{t+1}\downarrow\xi^R_{t+1}} R^b_{t+1}\left(\xi_{t+1},1\right) = R^b_{t+1}\left(\xi^R_{t+1},0\right) > \lim_{\xi_{t+1}\uparrow\xi^R_{t+1}} R^b_{t+1}\left(\xi_{t+1},1\right). \tag{73}$$

This implies that, if the capital quality shock is at the aggregate run threshold ξ^R_{t+1}, an increase in leverage from the value that makes the recovery rate equal to unity at liquidation prices, does not induce default as long as it is not so large that the bank becomes insolvent even at normal prices.

By definition of the run threshold ξ^R_{t+1}, the value of leverage that makes the recovery rate at liquidation prices equal to unity is exactly the aggregate leverage $\bar{\phi}_t$, that is

$$\frac{\bar{\phi}_t - 1}{\bar{\phi}_t}\overline{R}_{t+1}\left(\bar{\phi}_t\right) = \lim_{\xi_{t+1}\uparrow\xi^R_{t+1}} R^b_{t+1}\left(\xi_{t+1},1\right).$$

On the other hand, we let $\hat{\phi}_t$ denote the value above which the bank defaults at the aggregate run threshold ξ^R_{t+1} even at normal prices. This value satisfies

$$\frac{\hat{\phi}_t - 1}{\hat{\phi}_t}\overline{R}_{t+1}\left(\hat{\phi}_t\right) = R^b_{t+1}\left(\xi^R_{t+1},0\right)$$

and (73) implies that $\hat{\phi}_t > \bar{\phi}_t$.

For any value of leverage above the aggregate level $\bar{\phi}_t$ but below $\hat{\phi}_t$, when a sunspot is observed, the bank defaults if and only if a system wide run happens. That is $\xi^R_{t+1}\left(\phi\right)$ is insensitive to variation in individual bank's leverage in this region:

$$\xi^R_{t+1}\left(\phi\right) = \xi^R_{t+1}(\overline{\phi}_t)\ for\ \phi \in [\overline{\phi}_t, \widehat{\phi}_t].$$

For values of leverage above $\widehat{\phi}_t$ the bank is always insolvent even at non liquidation prices whenever it defaults, *i.e.* $\xi^R_{t+1}\left(\phi\right) = \xi^I_{t+1}\left(\phi\right)$ for $\phi > \hat{\phi}_t$. When ϕ is smaller than aggregate $\overline{\phi}_t$, the bank is less vulnerable to the run so that $\xi^R_{t+1}(\phi) < \xi^R_{t+1}$. In the extreme when the leverage multiple equals unity, the individual bank is not vulnerable to run so that $\xi^R_{t+1}(1) = 0$.

To summarize, the behavior of $\xi^R_{t+1}(\phi)$ can be characterized as follows:

$$\begin{gathered}\lim_{\phi\downarrow 1}\xi^R_{t+1}(\phi)=0\\ \frac{d\xi^R_{t+1}(\phi)}{d\phi}>0,\ \text{for } \phi\in\left(1,\overline{\phi}_t\right)\\ \xi^R_{t+1}(\phi)=\xi^R_{t+1},\ \text{for } \phi\in[\overline{\phi}_t,\widehat{\phi}_t] \text{ where } \xi^I_{t+1}(\widehat{\phi}_t)=\xi^R_{t+1}\\ \xi^R_{t+1}(\phi)=\xi^I_{t+1}(\phi),\ \text{for } \phi\in[\widehat{\phi}_t,\infty).\end{gathered} \tag{74}$$

See Figure A-1.

We can now rewrite the problem of the bank as in the text, but incorporating explicitly the dependence of the default and non default sets on the individual choice of leverage, as captured by $\Xi^D_{t+1}(\phi)$ and $\Xi^N_{t+1}(\phi)$:

$$\max_{\phi}\ \left(\mu_t\phi+\nu_t\right), \tag{75}$$

subject to the incentive constraint:

$$\theta\phi\leq\mu_t\phi+\nu_t, \tag{76}$$

the deposit rate constraint obtained from (67) :

$$\overline{R}_{t+1}(\phi)=\frac{\left[1-\frac{\phi}{\phi-1}\int_{\Xi^D_{t+1}(\phi)}\Lambda_{t+1}R^b_{t+1}(\xi_{t+1},\iota_{t+1})d\widetilde{F}_t\right]}{\int_{\Xi^N_{t+1}(\phi)}\Lambda_{t+1}d\widetilde{F}_t}. \tag{77}$$

μ_t and ν_t given by

$$\mu_t=\int_{\Xi^N_{t+1}(\phi)}\Omega_{t+1}[R^b_{t+1}-\overline{R}_{t+1}(\phi)]d\widetilde{F}_t \tag{78}$$

$$\nu_t=\int_{\Xi^N_{t+1}(\phi)}\Omega_{t+1}\overline{R}_{t+1}(\phi)d\widetilde{F}_t. \tag{79}$$

and where $\Xi^D_{t+1}(\phi)$ and $\Xi^N_{t+1}(\phi)$ are given by $(70)-(71)$, $\xi^I_{t+1}(\phi)$ and $\xi^R_{t+1}(\phi)$ satisfy $(68)-(69)$.

Using $(78)-(79)$ in the objective we can write the objective function as

$$\Psi_t(\phi)=\int_{\Xi^N_{t+1}(\phi)}\Omega_{t+1}\{[R^b_{t+1}(\xi_{t+1},\iota_{t+1})-\overline{R}_{t+1}(\phi)]\phi+\overline{R}_{t+1}(\phi)\}d\widetilde{F}_t. \tag{80}$$

Before proceeding with differentiation of the objective above, we introduce some notation that will be helpful in what follows. For any function $G\left(\xi_{t+1}, \iota_{t+1}\right)$ and for any ϕ different from $\bar{\phi}_t$ or $\hat{\phi}_t$ we let

$$
\begin{aligned}
(G)^*_{\phi t} &\equiv \frac{d}{d\phi}\left[\int_{\Xi^D_{t+1}(\phi)} G(\xi,\iota) d\widetilde{F}_t(\xi,\iota)\right] \\
&= \frac{d}{d\phi}\left[(1-\varkappa)\int_0^{\xi^I_{t+1}(\phi)} G(\xi,0)dF_t(\xi) + \varkappa\int_0^{\xi^R_{t+1}(\phi)} G(\xi,1)dF_t(\xi)\right]
\end{aligned}
$$

$$
= (1-\varkappa)G(\xi^I_{t+1}(\phi),0)f_t\left(\xi^I_{t+1}(\phi)\right)\frac{d\xi^I_{t+1}(\phi)}{d\phi} + \varkappa G(\xi^R_{t+1}(\phi),1)f_t\left(\xi^R_{t+1}(\phi)\right)\frac{d\xi^R_{t+1}(\phi)}{d\phi}.
$$

Then we know that as long as $G(\cdot)$ is continuous at $\xi^I_{t+1}(\phi)$ and $\xi^R_{t+1}(\phi)$ we have

$$
\frac{d}{d\phi}\left[\int_{\Xi^N_{t+1}(\phi)} G(\xi,\iota) d\widetilde{F}_t(\xi,\iota)\right] = -(G)^*_{\phi t}.
$$

Notice that we have not defined $(G)^*_{\phi t}$ for $\phi_t = \bar{\phi}_t$ or $\phi_t = \hat{\phi}_t$ because $\frac{d\xi^R_{t+1}(\phi)}{d\phi}$ does not exist at that point.

Differentiation of (80) at any value different from $\bar{\phi}_t$ and $\hat{\phi}_t$ yields

$$
\Psi'_t(\phi) = \mu_t - (\phi-1)\frac{\nu_t}{\overline{R}_{t+1}}\frac{d\overline{R}_{t+1}(\phi)}{d\phi} - \left(\Omega_{t+1}\left\{[R^b_{t+1}(\xi_{t+1},\iota_{t+1}) - \overline{R}_{t+1}(\phi)]\phi + \overline{R}_{t+1}(\phi)\right\}\right)^*_{\phi t}
$$

now notice that for $\phi \in [1, \bar{\phi}_t)$ and $\phi > \hat{\phi}_t$ we have that the bank networth is zero at both thresholds, that is

$$
\begin{aligned}
&[R^b_{t+1}(\xi^I_{t+1}(\phi),0) - \overline{R}_{t+1}(\phi)]\phi + \overline{R}_{t+1}(\phi) = 0 \\
&[R^b_{t+1}(\xi^R_{t+1}(\phi),1) - \overline{R}_{t+1}(\phi)]\phi + \overline{R}_{t+1}(\phi) = 0
\end{aligned}
$$

implying $\left(\Omega_{t+1}\left\{[R^b_{t+1}(\xi_{t+1},\iota_{t+1}) - \overline{R}_{t+1}(\phi)]\phi + \overline{R}_{t+1}(\phi)\right\}\right)^*_{\phi t} = 0$.

For $\phi \in \left(\bar{\phi}_t, \hat{\phi}_t\right)$ we have that at the insolvency threshold net worth is still zero

$$
[R^b_{t+1}(\xi^I_{t+1}(\phi),0) - \overline{R}_{t+1}(\phi)]\phi + \overline{R}_{t+1}(\phi) = 0
$$

while the run threshold is fixed at the aggregate level

$$
\frac{d\xi^R_{t+1}(\phi)}{d\phi} = 0
$$

so that again $\left(\Omega_{t+1}\left\{[R^b_{t+1}(\xi_{t+1},\iota_{t+1})-\overline{R}_{t+1}(\phi)]\phi+\overline{R}_{t+1}(\phi)\right\}\right)^*_{\phi t}=0.$

Therefore we have that for all ϕ different from $\bar{\phi}_t$ and $\hat{\phi}_t$

$$\Psi_t'(\phi)=\mu_t-(\phi-1)\frac{\nu_t}{\overline{R}_{t+1}}\frac{d\overline{R}_{t+1}(\phi)}{d\phi}$$

and by continuity of $\Psi_t(\phi)$ and $\mu_t-(\phi-1)\frac{\nu_t}{\overline{R}_{t+1}}\frac{d\overline{R}_{t+1}(\phi)}{d\phi}$ it can be extended to $\bar{\phi}_t$ and $\hat{\phi}_t$ as well.

Then, as reported in the text, the first order condition is

$$\begin{array}{c}\phi_t=\frac{\nu_t}{\theta-\mu_t},\ \text{if}\ \mu_t^r>0,\ \text{and}\\ \mu_t^r=0,\ \text{if}\ \phi_t<\frac{\nu_t}{\theta-\mu_t},\end{array}\tag{81}$$

$$\mu_t^r=\mu_t-(\phi_t-1)\frac{\nu_t}{\overline{R}_{t+1}}\frac{d\overline{R}_{t+1}(\phi_t)}{d\phi_t}\tag{82}$$

where(Here we assume $\mu_t<\theta$ which we will verify later).

As explained below, see section 6.5, we make assumptions such that conditions (81) – (82) characterize the unique global optimum for the bank's choice of leverage. Since these conditions don't depend on the individual net worth of a banker, every banker chooses the same leverage multiple and has the same Tobin's Q

$$\psi_t=\mu_t\phi_t+\nu_t.\tag{83}$$

Thus from the discussion in the text, it follows that there is a system wide default if and only if

$$R^b_{t+1}(\xi_{t+1},0)=\xi_{t+1}\frac{Z_{t+1}+(1-\delta)Q_{t+1}}{Q_t}<\frac{\phi_t-1}{\phi_t}\overline{R}_{t+1}(\phi_t),\ \text{or}$$

$$R^b_{t+1}(\xi_{t+1},1)=\xi_{t+1}\frac{Z_{t+1}+(1-\delta)Q^*_{t+1}}{Q_t}<\frac{\phi_t-1}{\phi_t}\overline{R}_{t+1}(\phi_t),$$

where $\overline{R}_{t+1}(\phi_t)$ is the aggregate promised deposit interest rate.

A systemic default occurs if and only if

$$\xi_{t+1}<\xi^I_{t+1},\ \text{where}\ \xi^I_{t+1}\frac{Z_{t+1}+(1-\delta)Q_{t+1}}{Q_t}=\frac{\phi_t-1}{\phi_t}\overline{R}_{t+1}(\phi_t),\tag{84}$$

or

$$\xi_{t+1} < \xi^R_{t+1} \text{ and } \iota_{t+1} = 1, \text{ where } \xi^R_{t+1} \frac{Z_{t+1} + (1-\delta)Q^*_{t+1}}{Q_t} = \frac{\phi_t - 1}{\phi_t} \overline{R}_{t+1}(\phi_t). \tag{85}$$

It follows that the probability of default at date t+1 conditional on date t information in the symmetric equilibrium is given by

$$p_t = F_t(\xi^I_{t+1}) + \varkappa \left[F_t(\xi^R_{t+1}) - F_t(\xi^I_{t+1}) \right]. \tag{86}$$

The aggregate capital holding of the banking sector is proportional to the aggregate net worth as

$$Q_t S^b_t = \phi_t N_t. \tag{87}$$

The aggregate net worth of banks evolves as

$$\begin{aligned} N_t &= \sigma \left\{ \xi_t \left[Z_t + (1-\delta) Q_t \right] S^b_{t-1} - \overline{R}_t D_{t-1} \right\} + \zeta S_{t-1}, \text{ if } \xi_t \left[Z_t + (1-\delta) Q_t \right] S^b_{t-1} \geq \overline{R}_t D_{t-1}, \\ N_t &= 0, \text{ otherwise.} \end{aligned} \tag{88}$$

Banks finance capital holdings by either net worth or deposit implies

$$D_t = (\phi_t - 1) N_t. \tag{89}$$

6.4 Market Clearing

The market for capital holding implies

$$S_t = S^b_t + S^h_t. \tag{90}$$

The final goods market clearing condition implies

$$Y_t = C_t + I_t + \frac{\rho^r}{2} {\pi_t}^2 Y_t + G. \tag{91}$$

As is explained in the text, the monetary policy rule is given by

$$R^n_t = \frac{1}{\beta} (\pi_t)^{\varphi_\pi} \left(\frac{MC_t}{\frac{\varepsilon - 1}{\varepsilon}} \right)^{\varphi_y}. \tag{92}$$

The recursive equilibrium is given by a set of ten quantity variables $(K_t, S_t, I_t, L_t, Y_t, C_t, S^h_t, S^b_t, D_t, N_t)$, seven price variables $(w_t, Z_t, MC_t, \pi_t, \overline{R}_{t+1},$

$Q_t, R^n_t)$ and six bank coefficients $\left(\psi_t, \mu_t, \nu_t, \mu^r_t, \phi_t, p_t, \xi^R_{t+1}, \xi^I_{t+1}\right)$ as a function of the four state variables $\mathcal{M}_t = (S_{t-1}, S^b_{t-1}, \overline{R}_t D_{t-1}, \xi_t)$ and a sunspot variable ι_t, which satisfies twenty three equations, given by: (54,55,56,59,60,61,62,63,64,65, 68,69,66,78,79,81,82,83,86,87,88,89,90,91,92). Here, the capital quality shocks follows a Markov process $\xi_{t+1} \sim F\left(\xi_{t+1} \mid \xi_t\right)$ and the sunspot is iid. with $\iota_t = 1$ with probability $\varkappa$.

6.5 On the Global Optimum for Individual Bank's Choice

To study global optimality of the individual leverage choice selected by the first order conditions in (81) we need to analyze the curvature of the objective function $\Psi_t(\phi)$ in (80).

To do so we use (77) to derive an expression for $\frac{d\bar{R}}{d\phi}$ and substitute it into (82) to obtain

$$\Psi'_t(\phi) = \int_{\Xi^N_{t+1}(\phi)} \Omega_{t+1} \left\{ R^b_{t+1} - \frac{\left[1 - \int_{\Xi^D_{t+1}(\phi)} \Lambda_{t+1} R^b_{t+1} d\widetilde{F}_t\right]}{\int_{\Xi^N_{t+1}(\phi)} \Lambda_{t+1} d\widetilde{F}_t} \right\} d\widetilde{F}_t \qquad (93)$$

To understand this expression, consider a banker that borrows an extra unit of the consumption good and uses it to purchase capital. The first term, R^b_{t+1}, captures the value to the bank of an increase in bank assets which the bank enjoys only if it does not default in the subsequent period. The second term is the marginal cost of borrowing one unit of the consumption good from depositors. When the bank defaults, the bank asset returns belong to the depositors which they value as $\int_{\Xi^D_{t+1}(\phi)} \Lambda_{t+1} R^b_{t+1} d\widetilde{F}_t$. When the bank does not default it pays a promised amount, $\bar{r}$, which depositors value as $\bar{r} \int_{\Xi^N_{t+1}(\phi)} \Lambda_{t+1} d\widetilde{F}_t$. Therefore, the required promised payment to depositors, $\bar{r}$, that makes them indifferent between consuming and lending this *additional* unit to the bank must satisfy

$$\int_{\Xi^D_{t+1}(\phi)} \Lambda_{t+1} R^b_{t+1} d\widetilde{F}_t + \bar{r} \int_{\Xi^N_{t+1}(\phi)} \Lambda_{t+1} d\widetilde{F}_t = 1.$$

Rearranging gives

$$\bar{r}(\phi) = \frac{1 - \int_{\Xi^D_{t+1}(\phi)} \Lambda_{t+1} R^b_{t+1} d\widetilde{F}_t}{\int_{\Xi^N_{t+1}(\phi)} \Lambda_{t+1} d\widetilde{F}_t}, \qquad (94)$$

that is the second term.

Notice that the marginal cost of deposits $\bar{r}(\phi)$ is higher than the average cost $\bar{R}_{t+1}(\phi)$ in equation (77). This is because the average cost also factors in that asset purchases are partially financed by bank net worth. In fact, using (77) and (94) we get

$$\bar{r}(\phi) = \frac{\phi - 1}{\phi}\overline{R}_{t+1}(\phi) + \frac{1}{\phi}\frac{1}{\int_{\Xi_{t+1}^{N}(\phi)} \Lambda_{t+1} d\widetilde{F}_t} \tag{95}$$

which says that the marginal cost of debt is a weighted average between the average cost $\overline{R}_{t+1}(\phi)$ and the required rate on a security that only pays in the no default region, $\frac{1}{\int_{\Xi_{t+1}^{N}(\phi)} \Lambda_{t+1} d\widetilde{F}_t}$, where the weights are given by the proportion of asset that is debt financed and equity financed respectively.

Inspecting the expression for $\Psi_t'(\phi_t)$ in (93) which we rewrite below as

$$\Psi_t'(\phi) = \int_{\Xi_{t+1}^{N}(\phi)} \Omega_{t+1} R_{t+1}^{b} d\widetilde{F}_t - \left[1 - \int_{\Xi_{t+1}^{D}(\phi)} \Lambda_{t+1} R_{t+1}^{b} d\widetilde{F}_t\right] \frac{\int_{\Xi_{t+1}^{N}(\phi)} \Omega_{t+1} R_{t+1}^{b} d\widetilde{F}_t}{\int_{\Xi_{t+1}^{N}(\phi)} \Lambda_{t+1} d\widetilde{F}_t}, \tag{96}$$

reveals that the objective of the banker is a non linear function of individual bank's leverage. In particular the individual choice of leverage of a banker affects its own probability of default, *i.e.* the set of states $\Xi_{t+1}^{D}(\phi)$ changes as ϕ changes. This in turn affects the marginal benefits and costs of increasing leverage as shown in equation (96).

Then proceeding as in section 6.3 to differentiate (96) for any value of ϕ different from $\bar{\phi}_t$ and $\hat{\phi}_t$, we get

$$\Psi_t"(\phi) = \begin{array}{c} -\left(\Omega_{t+1} R_{t+1}^{b}\right)_{\phi t}^{*} + \left(\Lambda_{t+1} R_{t+1}^{b}\right)_{\phi t}^{*} \cdot \frac{\int_{\Xi_{t+1}^{N}(\phi)} \Omega_{t+1} d\widetilde{F}_t}{\int_{\Xi_{t+1}^{N}(\phi)} \Lambda_{t+1} d\widetilde{F}_t} \\ + \\ \bar{r}(\phi) \int_{\Xi_{t+1}^{N}(\phi)} \Omega_{t+1} R_{t+1}^{b} d\widetilde{F}_t \left[\frac{\left(\Omega_{t+1}\right)_{\phi t}^{*}}{\int_{\Xi_{t+1}^{N}(\phi)} \Omega_{t+1} d\widetilde{F}_t} - \frac{\left(\Lambda_{t+1}\right)_{\phi t}^{*}}{\int_{\Xi_{t+1}^{N}(\phi)} \Lambda_{t+1} d\widetilde{F}_t}\right] \end{array} \tag{97}$$

Note that for $\phi \in [1, \bar{\phi}_t)$

$$R_{t+1}^{b}\left(\xi_{t+1}^{I}(\phi), \iota\right) = R_{t+1}^{b}\left(\xi_{t+1}^{R}(\phi), 1\right) = \frac{\phi - 1}{\phi}\overline{R}_{t+1}(\phi)$$

For $\phi \in \left(\bar{\phi}_t, \hat{\phi}_t\right)$ we have $\frac{d\xi^R_{t+1}(\phi)}{d\phi} = 0$ which implies that for any function $G\left(\xi_{t+1}, \iota_{t+1}\right)$

$$(G)^*_{\phi t} = (1-\varkappa)G(\xi^I_{t+1}(\phi),0)f_t\left(\xi^I_{t+1}(\phi)\right)\frac{d\xi^I_{t+1}(\phi)}{d\phi} \text{ for } \phi \in \left(\bar{\phi}_t, \hat{\phi}_t\right) \quad (98)$$

and also

$$R^b_{t+1}\left(\xi^I_{t+1}(\phi),0\right) = \frac{\phi-1}{\phi}\overline{R}_{t+1}(\phi)$$

Then, we learn

$$\begin{aligned}\left(\Omega_{t+1}R^b_{t+1}\right)^*_{\phi t} &= (\Omega_{t+1})^*_{\phi t}\cdot\frac{\phi-1}{\phi}\overline{R}_{t+1}(\phi)\\ \left(\Lambda_{t+1}R^b_{t+1}\right)^*_{\phi t} &= (\Lambda_{t+1})^*_{\phi t}\cdot\frac{\phi-1}{\phi}\overline{R}_{t+1}(\phi).\end{aligned}$$

Substituting this back into (97) we get

$$\Psi_t\text{''}(\phi) = \left(\bar{r}(\phi) - \frac{\phi-1}{\phi}\overline{R}_{t+1}(\phi)\right)\int_{\Xi^N_{t+1}(\phi)}\Omega_{t+1}d\widetilde{F}_t\left[\frac{(\Omega_{t+1})^*_{\phi t}}{\int_{\Xi^N_{t+1}(\phi)}\Omega_{t+1}d\widetilde{F}_t} - \frac{(\Lambda_{t+1})^*_{\phi t}}{\int_{\Xi^N_{t+1}(\phi)}\Lambda_{t+1}d\widetilde{F}_t}\right]$$

Finally, from (94), we also notice

$$\bar{r}(\phi) = \frac{\phi-1}{\phi}\overline{R}_{t+1}(\phi) + \frac{1}{\phi\int_{\Xi^N_{t+1}(\phi)}\Lambda_{t+1}d\widetilde{F}_t}.$$

Therefore, we get

$$\Psi_t\text{''}(\phi) = \frac{1}{\phi}\frac{\int_{\Xi^N_{t+1}(\phi)}\Omega_{t+1}d\widetilde{F}_t}{\int_{\Xi^N_{t+1}(\phi)}\Lambda_{t+1}d\widetilde{F}_t}\left[\frac{(\Omega_{t+1})^*_{\phi t}}{\int_{\Xi^N_{t+1}(\phi)}\Omega_{t+1}d\widetilde{F}_t} - \frac{(\Lambda_{t+1})^*_{\phi t}}{\int_{\Xi^N_{t+1}(\phi)}\Lambda_{t+1}d\widetilde{F}_t}\right] \quad (99)$$

for any ϕ different from $\bar{\phi}_t$ and $\hat{\phi}_t$.[21]

[21] Notice that $(\Omega_{t+1})^*_{\phi t}$ and $(\Lambda_{t+1})^*_{\phi t}$ are not continuous at $\bar{\phi}_t$ since, for instance

$$\begin{aligned}\lim_{\phi\uparrow\bar{\phi}_t}(\Omega_{t+1})^*_{\phi t} &= (1-\varkappa)\Omega_{t+1}(\xi^I_{t+1},0)f_t\left(\xi^I_{t+1}\right)\frac{d\xi^I_{t+1}(\phi)}{d\phi} + \varkappa\Omega_{t+1}(\xi^R_{t+1},1)f_t\left(\xi^R_{t+1}\right)\frac{d^L\xi^R_{t+1}(\phi)}{d\phi} > \\ &> (1-\varkappa)\Omega_{t+1}(\xi^I_{t+1},0)f_t\left(\xi^I_{t+1}\right)\frac{d\xi^I_{t+1}(\phi)}{d\phi} = \lim_{\phi\downarrow\bar{\phi}_t}(\Omega_{t+1})^*_{\phi t}\end{aligned}$$

where $\frac{d^L\xi^R_{t+1}(\phi)}{d\phi}$ is the left derivative of $\xi^R_{t+1}(\phi)$ at $\bar{\phi}_t$. This implies that $\Psi\text{''}(\phi)$ does not exist at $\bar{\phi}_t$.

From equations (98) and (99) we learn that , above the aggregate level of leverage $\bar{\phi}_t$, the curvature of the objective function of the bank depends on the relative increase in the marginal value of wealth of the banker and of the households *only at the insolvency threshold*. This region however is only relevant in the case in which the constraint is not binding and in this case, in our calibration, the probability of default is always zero, so that the objective is flat to the right of $\bar{\phi}_t$. This implies that we can study optimality by only looking at the region below the aggregate level $\bar{\phi}_t$.

If individual leverage is below the aggregate level of leverage $\bar{\phi}_t$, the curvature of the objective function of the bank depends on the relative increase in the marginal value of wealth of the banker and of the households *both at the insolvency threshold and at the run threshold*. Quantitatively the probability of insolvency is much smaller than the probability of a run so that, in this region, the curvature is mainly determined by what we assume about $\Omega_{t+1}(\xi^R_{t+1}(\phi), 1)$: the marginal value of wealth of a bank that decreases its own leverage below the aggregate value $\bar{\phi}_t$ and survives a run at a threshold $\xi^R_{t+1}(\phi) < \xi^R_{t+1}$. We assume that a bank that survives a systemic bank run behaves just like new entrants during the panic: it stores its net worth and starts operating the period right after the crisis. Given that both leverage and spreads increase dramatically afer a crisis, new banker's Tobin's Q is very high during a crisis so that (99) implies that the objective function of the banker is strictly convex in the region where leverage is below the aggregate level $\bar{\phi}_t$, that is $\Psi_t"(\phi) > 0$ for $\phi \in [1, \bar{\phi}_t)$. Therefore, global optimality can be checked by comparing the equilibrium franchise value of the bank, associated with the equilibrium choice $\phi = \bar{\phi}_t$, to the value of deviating to the corner where leverage is unity, i.e. $\phi = 1$.

Even if it was costless for a bank to decrease its leverage to unity, the deviation would still not be profitable around the risk adjusted steady state. However it would become profitable when either the constraint is slack or the probability of the run increases enough. The only equilibrium in these states would then be one in which a proportion of banks decrease their leverage in anticipation of a run while all of the others are against the constraint, i.e. there is no symmetric equilibrium. However, we think that there are costs associated to the drastic downsizing of operations associated to a deviation of this kind. What we have in mind are reputations costs associated with the bank's refusal to accept deposits in a given period in order to survive a run in the subsequent period. In particular, we posit that the objective of the bank is given by

$$V_t(n_t) = \Psi_t(\phi)\, n_t \left(1 - \tau\bar{\phi}_t\right) \quad for\ \phi \in [1, \bar{\phi}_t).$$

That is, a deviation of a bank that reduces leverage below the aggregate value $\bar{\phi}_t$ entails a fixed cost $\tau\bar{\phi}_t$ per unit of net worth. We check computationally that the deviation is never profitable, i.e. $\Psi_t\left(\bar{\phi}_t\right) > \Psi_t(1)$, in all of our experiments for values of τ which are greater than or equal to .77%.[22]

6.6 Computation

It is convenient for computations to let the aggregate state of the economy be given by

$$\mathcal{M}_t = (S_{t-1}, N_t, \xi_t, \iota_t).$$

Notice that bank net worth replaces the specific asset and liability position of banks in the natural state that we have used so far $\tilde{\mathcal{M}}_t = (S_{t-1}, S^b_{t-1}, D_{t-1}\bar{R}_t, \xi_t)$. To see that this state is sufficient to compute the equilibrium we rewrite the evolution of net worth, equation (88), forward. Using the definition of the leverage multiple and the budget constraint of the banker we get

$$N_{t+1} = \begin{cases} \sigma N_t \left\{\phi_t \left(\xi_{t+1} \frac{Z_{t+1} + (1-\delta) Q_{t+1}}{Q_t} - \overline{R}_{t+1}\right) + \overline{R}_t\right\} + \zeta S_t & \begin{array}{c} \text{if } \xi_{t+1} < \xi^I_{t+1} \\ \text{or } \xi_{t+1} < \xi^R_{t+1} \text{ and } \iota_{t+1} = 1 \end{array} \\ 0 & Otherwise. \end{cases} \tag{100}$$

We can then look for a recursive equilibrium in which each equilibrium variable is a function of $\mathcal{M}_t$ and the evolution of net worth is given by a function $N_{t+1}\left(\mathcal{M}_t;\ \xi_{t+1}, \iota_{t+1}\right)$ that depends on the realization of the exogenous shocks $\left(\xi_{t+1}, \iota_{t+1}\right)$ and satisfies equation (100) above.

We use time iteration in order to approximate the functions

$$\boldsymbol{\vartheta} = \left\{\mathbf{Q}(\mathcal{M}), \mathbf{C}(\mathcal{M}), \boldsymbol{\psi}(\mathcal{M}), \boldsymbol{\xi}^R_{t+1}(\mathcal{M}), \boldsymbol{\xi}^I_{t+1}(\mathcal{M}), \mathbf{T}(\mathcal{M}; \xi', \iota')\right\}$$

where $\mathbf{T}(\mathcal{M}_t; \xi', \iota')$ is the transition law determining the stochastic evolution of the state.

The computational algorithm proceeds as follows:

[22] The value of deviating can increase in very extreme cases but in a simulation of 100 thoushands periods it is still below 1.7% for 99 percent of the times.

1. Determine a functional space to use for approximating equilibrium functions. (We use piecewise linear).

2. Fix a grid of values for the state $G \subset \left[S^m, S^M\right] \times \left[0, N^M\right] \times \left[1 - 4\sigma^\xi, 1 + 4\sigma^\xi\right] \times \{0, 1\}$

3. Set $j = 0$ and guess initial values for the equilibrium objects of interest on the grid

$$\vartheta_j = \left\{Q_j\left(\mathcal{M}\right), C_j\left(\mathcal{M}\right), \psi_j\left(\mathcal{M}\right), \xi^R_{t+1,j}\left(\mathcal{M}\right), \xi^I_{t+1,j}\left(\mathcal{M}\right), \mathcal{T}_j\left(\mathcal{M}; \xi', \iota'\right)\right\}_{\mathcal{M} \in G}$$

4. Assume that ϑ_i has been found for $i < M$ where M is set to 10000. Use ϑ_i to find associated functions $\boldsymbol{\vartheta}_i$ in the approximating space, e.g. $\mathbf{Q}_i$ is the price function that satisfies $\mathbf{Q}_i\left(\mathcal{M}\right) = Q_i\left(\mathcal{M}\right)$ for each $\mathcal{M} \in G$.

5. Compute all time $t+1$ variables in the system of equilibrium equations by using the functions $\boldsymbol{\vartheta}_i$ from the previous step, e.g. for each $\mathcal{M} \in G$ let $Q_{t+1} = \mathbf{Q}_i\left(\mathcal{T}_j\left(\mathcal{M}; \xi', \iota'\right)\right)$, and then solve the system to get the implied ϑ_{i+1}

6. Repeat 4 and 5 until convergence of ϑ_i