Booms and Systemic Banking Crises*

Frédéric Boissay† — European Central Bank
Fabrice Collard‡ — University of Bern
Frank Smets§ — European Central Bank

This draft: January 2, 2013
First draft: December 2011

Abstract

The empirical literature on systemic banking crises (SBCs) has shown that SBCs are rare events that break out in the midst of credit intensive booms and bring about particularly deep and long–lasting recessions. We attempt to explain these phenomena within a dynamic general equilibrium model featuring a non–trivial banking sector. In the model, banks are heterogeneous with respect to their intermediation skills, which gives rise to an interbank market. Moral hazard and asymmetric information on this market may generate sudden interbank market freezes, SBCs, credit crunches and, ultimately, severe recessions. Simulations of a calibrated version of the model indicate that typical SBCs break out in the midst of a credit boom generated by a sequence of positive supply shocks rather than being the outcome of a big negative wealth shock. We also show that the model can account for the relative severity of recessions with SBCs and their longer duration.

Keywords: Moral Hazard, Asymmetric Information, Lending Boom, Credit Crunch, Systemic Banking Crisis

JEL Class.: E32, E44, G01, G21.

*We benefited from discussions with F. Alvarez, F. Canova, S. Claessens, R. Cooper, V. Curdìa, M. Darracq-Paries, F. Dufourt, H. Degryse, X. Freixas, S. Gilchrist, J. Gomes, P–O. Gourinchas, J. Henry, I. Jaccard, P. Jacquinot, P. Karadi, N. Kiyotaki, L. Laeven, G. Lombardo, K. Nikolov, G. Nuño, H–S. Shin, J. Stein, G. Ströbl, J. Suarez, M. Trabandt, P. Weil, as well as seminar participants at the ECB, the European University Institute, the University of Munich, CRETE 2012, CEPR conference in St Gallen, Banque de France Conference in Strasbourg, National Bank of Belgium, REDg 2012 Workshop in Madrid, Conference in honor of C. Sims in Princeton, Bocconi University, 3rd Joint French macro workshop in Paris, National Bank of Serbia, University of Bergen. We are particularly indebted to H. Dellas and B. Diba.

†European Central Bank, Postfach 16 03 19, 60066 Frankfurt am Main, Germany. email: frederic.boissay@ecb.int, URL:`http://www.fboissay.eu`

‡Universität Bern, Department of Economics, Schanzeneckstrasse 1, Postfach 8573, CH-3001 Bern. email: fabrice.collard@gmail.com, URL:`http://fabcol.free.fr`

§European Central Bank, Postfach 16 03 19, 60066 Frankfurt am Main, Germany. email: frank.smets@ecb.int

Non–Technical Summary

Recent empirical research on systemic banking crises (henceforth, SBCs) has highlighted the existence of similar patterns across diverse episodes. SBCs are rare events. Recessions that follow SBC episodes are deeper and longer lasting than other recessions. And, more importantly for the purpose of this paper, SBCs follow credit intensive booms; "banking crises are credit booms gone wrong" (Schularick and Taylor, 2012, p. 1032). Rare, large, adverse financial shocks could possibly account for the first two properties. But they do not seem in line with the fact that the occurrence of an SBC is not random but rather closely linked to credit conditions. So, while most of the existing macro–economic literature on financial crises has focused on understanding and modeling the propagation and the amplification of adverse random shocks, the presence of the third stylized fact mentioned above calls for an alternative approach.

In this paper we develop a simple macroeconomic model that accounts for the above three stylized facts. The primary cause of systemic banking crises in the model is the accumulation of assets by households in anticipation of future adverse shocks. The typical run of events leading to a financial crisis is as follows. A sequence of favorable, non permanent, supply shocks hits the economy. The resulting increase in the productivity of capital leads to a demand–driven expansion of credit that pushes the corporate loan rate above steady state. As productivity goes back to trend, firms reduce their demand for credit, whereas households continue to accumulate assets, thus feeding the supply of credit by banks. The credit boom then turns supply–driven and the corporate loan rate goes down, falling below steady state. By giving banks incentives to take more risks or misbehave, too low a corporate loan rate contributes to erode trust within the banking sector precisely at a time when banks increase in size. Ultimately, the credit boom lowers the resilience of the banking sector to shocks, making systemic crises more likely.

We calibrate the model on the business cycles in the US (post WWII) and the financial cycles in fourteen OECD countries (1870–2008), and assess its quantitative properties. The model reproduces the stylized facts associated with SBCs remarkably well. Most of the time the model behaves like a standard financial accelerator model, but once in while —on average every forty years— there is a banking crisis. The larger the credit boom, (i) the higher the probability of an SBC, (ii) the sooner the SBC, and (iii) —once the SBC breaks out— the deeper and the longer the recession. In our simulations, the recessions associated with SBCs are significantly deeper (with a 45% larger output loss) than average recessions. Overall, our results validate the role of supply–driven credit booms leading to credit busts. This result is of particular importance from a policy making perspective as it implies that systemic banking crises are predictable. We indeed use the model to compute the k–step ahead probability of an SBC at any point in time. Fed with actual US data over the period 1960–2011, the model yields remarkably realistic results. For example, the one–year ahead probability of a crisis is essentially zero in the 60–70s. It jumps up twice during the sample period: in 1982–3, just before the Savings & Loans crisis, and in 2007–9. Although very stylized, our model thus also provides with a simple tool to detect financial imbalances and predict future crises.

1 Introduction

Recent empirical research on systemic banking crises (henceforth, SBCs) has highlighted the existence of similar patterns across diverse episodes (see Reinhart and Rogoff, 2009; Jordà et al., 2011a,b; Claessens et al., 2011; Schularick and Taylor, 2012). SBCs are rare events. Recessions that follow SBC episodes are deeper and longer lasting than other recessions (see Section 2). And, more importantly for the purpose of this paper, SBCs follow credit intensive booms; "banking crises are credit booms gone wrong" (Schularick and Taylor, 2012, Borio and Drehmann, 2009, and Borio and Lowe, 2002; the notion that banking crises are endogenous and follow prosperous times is also present in Minsky, 1977). Most of the existing macro–economic literature on financial crises has focused on understanding and modeling the propagation and the amplification of random adverse shocks. Indeed, rare, large enough, adverse financial shocks can account for the first two properties (see *e.g.* Gertler and Kiyotaki, 2009). However, by implying that financial crises may break out at any time in the business cycle, they do not seem in line with the fact that the occurrence of an SBC is closely linked to credit conditions (Gorton, 2010, 2012). The third stylized fact therefore calls for an alternative approach.

In this paper, financial crises result from the pro–cyclicality of bank balance sheets that emanates from interbank market funding. During expansions, bank market funding and credit supply increase, pushing down the rates of return on corporate and interbank loans. The lower rates accentuate agency problems in the interbank market that lead to a reduction on market funding and contractions. The larger the credit boom relative to the possibilities for productive use of loans, the larger the fall in interest rates, and the higher the probability of a bank run in —and therefore of a disastrous freeze of— the interbank market.[1] As in Shin (2008) and Hahm et al. (2011), the behavior of banks (credit in our case) during good times sows the seeds of a financial crisis. In our model, banks are heterogeneous in terms of —non–publicly observed— intermediation efficiency. They finance their activities with funds obtained from depositors/shareholders or raised in the interbank market. There exists the usual agency problem in this market as borrowing banks can always divert some of the funds into low return assets that cannot be recovered by the lending banks. The incentives for diversion are stronger for less productive banks and depend on the level of interest rates in the economy. The lower the return on loans, the greater the incentive to engage in fund diversion and hence the greater counterparty risk in the interbank market. The typical run of events leading to a financial crisis is as follows. A sequence of favorable,

[1] Our representation of financial crises as market–based bank runs is in line with what happened during the 2007-8 financial crisis (see Uhlig, 2010). Shin (2010, Chap. 8), for example, depicts the demise of Northern Rock —a UK bank— in 2007 as primarily originating from the sudden freezing of the short–term funding market, what he refers to as a "modern bank run". A traditional, deposit–based, run on the bank took place as well, but it did so one month later, accounted for only 10% of the bank's fall in total funding, and rapidly stopped because, following the news of the run, the UK authorities pledged 100% deposit guarantees.

non permanent, supply shocks hits the economy. The resulting increase in the productivity of capital leads to a demand–driven expansion of credit that pushes interest rates up. The more efficient banks expand their loan operations by drawing funds from the less efficient banks and market funding in the banking sector as a whole increases. The economy booms. But as the supply shocks run their course, the probability of imminent reversion to average productivity increases. This slows down corporate demand for loans while at the same time inducing households to accumulate savings in order to smooth consumption. Credit expansion becomes supply–driven, putting downward pressure on interest rates. The rate of return on interbank loans declines making the less efficient banks more prone to borrow themselves and divert those funds. As the identity of these banks is not known, counterparty risk in the interbank market goes up, interbank loans decline, and market finance recedes. The stronger the credit expansion during the booming times, the larger the decline in interest rates and the more acute the agency problem in the interbank market. We show that there is a threshold value of interest rates below which the interbank market freezes, corporate credit collapses and the economy tanks. This threshold can be alternatively expressed in terms of the level of banking assets relative to the level of productivity (output) in the economy, that we call the absorption capacity of the banking sector. Supply–driven, excessive credit creation places the economy beyond its absorption capacity, triggering an SBC.

Our work differs from related work on financial crises in several important aspects. In contrast to Shin (2008) and Hahm et al. (2011), whose models are static, ours is a full blown dynamic stochastic general equilibrium (DSGE) model, and thus more suitable for quantitative analysis. Unlike Bernanke et al. (1999), Jermann and Quadrini (2010), Gertler and Karadi (2011), who study the linearized system dynamics around the steady state in models where adverse shocks are amplified by financial market frictions our model analysis characterizes the full equilibrium dynamics inclusive of important and critical non-linearities such as the freezing of interbank markets. This is an important difference because near the steady state our model features a traditional financial accelerator. But away from it (and the large departures from the steady state are the endogenous outcome of a boom-bust endogenous cycle, rather than a big shock) it gives rise to banking crises. Crises are rare but generate particularly large output losses and inefficiencies due to the presence of pecuniary externalities. The models of Bianchi (2009), Bianchi and Mendoza (2010), and Korinek (2010) also exhibit non–linearities and pecuniary externalities but assume that the interest rate is exogenous, so they are at best applicable to small open economies and emerging markets.[2] Perhaps the models the closest to ours are Brunnermeier and Sannikov (2012) and He and Krishnamurthy (2012). These two models too feature a powerful non–linear amplification mechanism. As in other

[2]In these models non–linearities are due to occasionally binding constraints, whereas in our case they are due to the economy switching from normal to crisis times. Gertler and Kiyotaki (2012) also develop a model with bank runs and regime switches. While in their model bank runs are deposit–based and unexpected, in ours they are market–based and, more importantly, agents are fully rational: they perfectly know and take into account the probability that runs will occur in the future.

DSGE models with financial frictions, financial crises are the outcomes of adverse exogenous financial shocks (*e.g.* to banks' net worth), whose size ultimately determines the size of the crises. In our case, in contrast, shocks play only a secondary role because crises are related to whether or not financial imbalances (*e.g.* credit boom, ballooning bank balance sheets) have built up in the first place. That is, crises may break out endogenously, even in the absence of negative shocks. Another important feature of our model is that it does not rely on financial shocks to generate banking crises; the technology shock is indeed the only exogenous source of uncertainty.[3]

We calibrate the model on the business cycles in the US (post–WWII) and the financial cycles in fourteen OECD countries (1870-2008), and assess its quantitative properties. The model reproduces the stylized facts associated with SBCs remarkably well. Most of the time bank assets remain below the threshold for financial crises and the model behaves like a standard financial accelerator model. But once in a while —on average every forty years— there is a banking crisis. The typical banking crisis in our simulations is preceded with a credit boom and brings about both a credit crunch and a recession. On the brink of an SBC, risk averse households accumulate precautionary savings and inadvertently fuel a credit boom, which brings credit creation even closer to the economy's absorption capacity. Our findings are in line with empirical evidence (see Schularick and Taylor, 2011, among others) and validate the role of supply–driven credit booms leading to credit busts. The larger the credit boom, *(i)* the higher the probability of an SBC, *(ii)* the sooner the SBC, and *(iii)* —once the SBC breaks out— the deeper and the longer the recession. In our simulations, a recession associated with SBCs is significantly deeper (with a 45% larger output loss) than the average recession. We use the model to compute the k–step ahead probability of an SBC at any point in time. Fed with actual US data of total factor productivity over the period 1960-2011, the model produces remarkably realistic results. For example, the one–year ahead probability of a crisis is essentially zero in the 60–70s. It jumps up twice during the sample period: in 1982–3, just before the Savings & Loans crisis, and in 2007–9.

The paper proceeds as follows. Section 2 briefly documents key empirical facts about the dynamics of systemic banking crises in 14 OECD countries for the period 1870–2008. Section 3 describes our theoretical framework, the micro-foundations of interbank market freezes and the dynamic implications of such events. Section 4 discusses our calibration strategy and presents our solution method. Section 5 analyses the quantitative implications of the model as well as its performance against the facts documented in Section 2. A last section concludes.

[3]Another difference with many existing models concerns the modeling of the financial friction. In Bianchi's model, for example, the friction affects the firms and operates through excess credit demand ("over-borrowing"), whereas in our model it operates through excess credit supply. Our model also differs from Brunnermeier and Sannikov's in that it is a discrete time model that we calibrate on US and OECD data, which allows us to confront our model with the data.

2 Key Facts on Systemic Banking Crises

Reinhart and Rogoff (2009), Claessens et al. (2011), Jordà et al. (2011a,b), and Schularick and Taylor (2012) recently documented that SBCs share, despite their variety, a few common regularities. Building upon this earlier work, we briefly describe in this section the key facts on SBCs, against which we will later assess the quantitative properties of our model. To do so we use the historical dataset assembled by Jordà et al. (2011a). This dataset comprises yearly observations for real GDP per capita, total domestic currency loans of banks and banking institutions to non–financial companies and households, banks' total assets, the dates of business cycle peaks, and the dates of banking crises from 1870 to 2008 for 14 OECD developed countries.[4] A banking crisis is defined as an event during which the financial sector experiences bank runs, sharp increases in default rates accompanied by large losses of capital that result in public intervention, bankruptcy, or the forced merger of major financial institutions (see Laeven and Valencia, 2008). For the purpose of the present paper, we further define as *systemic* those banking crises that are concomitant with a recession, *i.e.* that break out between the peak and the trough of a given business cycle. Jordà et al. use the Bry-Boschan algorithm to date peaks and troughs consistently across countries. We exclude war times and only keep complete business cycles, from peak to peak. After trimming, our sample covers 176 full–length business cycles. The main statistics are reported in Table 1.

Table 1: Statistics on recessions and banking crises

	N. obs.	N	Frequency (%)	Magnitude (%) from peak to trough	Duration (Years) from peak to trough
All banking crises	1,736	78	4.49	–	–
Systemic Banking Crises (SBC)	1,736	42	2.42	–	–
All recessions	1,736	176	10.20	4.86 (5.91)	1.85
Recessions with SBC (A)	1,736	42	23.86	6.74 (6.61)	2.59
Recessions w/o SBC (B)	1,736	134	76.13	4.27 (5.61)	1.61
Test A≠B, p-value (%)	–	–	–	2.61	0.00

Note: the magnitudes reported into parentheses are calculated using the HP–filtered series of output, and are thereby corrected for the underlying trend in output. Following Ravn and Uhlig (2002), we set the parameter of the Hodrick–Prescott filter to 6.25.

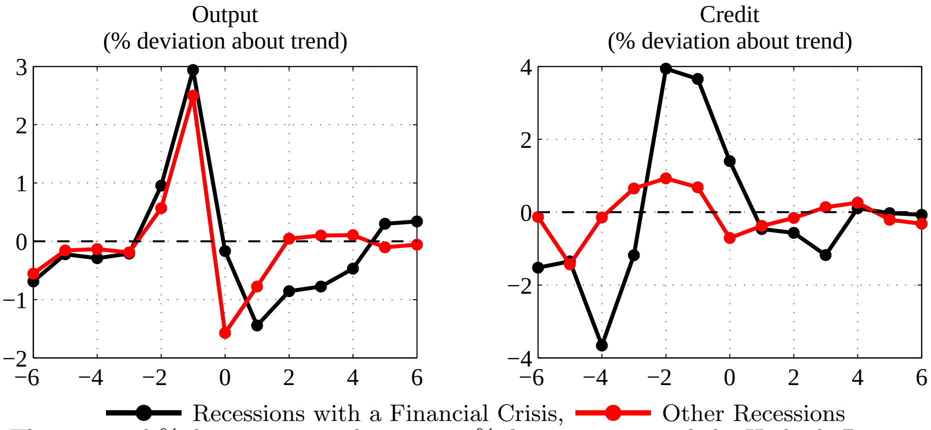
Fact #1: Systemic Banking Crises are Rare Events. 78 banking crises can be identified in the sample, which comprises 1,736 observations. The frequency of crises is therefore 4.49%, which means that countries in our sample experience a crisis, on average, every 22 years. Half of those 78 banking crises were systemic. Hence, SBCs are rare events, which occur on average every forty years. In contrast, recessions are much more frequent and occur

[4] The list of country comprises Australia, Canada, Denmark, France, Germany, Great Britain, Italy, Japan, the Netherlands, Norway, Spain, Sweden, Switzerland and the United States of America.

every ten years or so.

Fact #2: Financial recessions are deeper and last longer than other recessions. While only one fourth of the recessions we identify involve a banking crisis, these "financial recessions" are on average significantly deeper than other, regular recessions. For instance, we find that the drop in real GDP per capita from peak to trough is 40% bigger during financial recessions (6.74%) than during the average recession (4.86%) (60% deeper than recessions without SBCs), or about 12% when the data are HP–filtered (see Table 1). On average, systemic banking crises also last one year longer. The dynamics of financial recessions too is different: they tend to be preceded by a faster increase in GDP and credit compared with other recessions, as Figure C.6 shows. Claessens et al. (2011) report similar patterns based on a shorter data set that includes emerging countries.[5]

Figure 1: Financial versus normal recessions



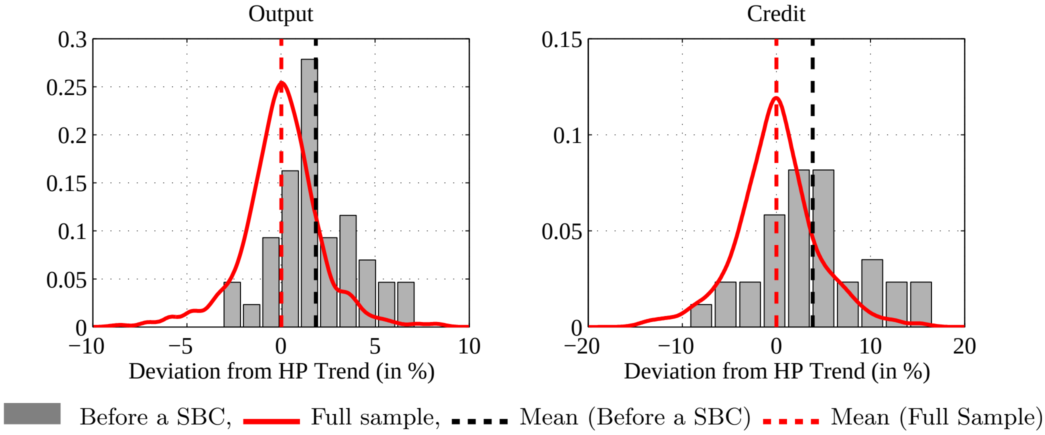
Note: The reported % deviations are the average % deviations around the Hodrick–Prescott trend (calculated with a parameter of 6.25). Notice that the implied magnitude of financial recessions in the left chart is about 4.3%, which is lower than that 6.61% reported earlier in Table 1. This discrepancy reflects the fact that the statistics in Table 1 also take into account recessions of more than 6 years. We find similar results when we consider the % deviations of output and credit from their respective linear trends (see the companion technical appendix).

Fact #3: Systemic banking crises break out in the midst of credit intensive booms. Systemic banking crises do not hit at random (Gorton, 1988). To illustrate this point, Figure 2 reports the empirical distributions of GDP (left panel) and credit (right panel) gaps, as measured by the percentage deviations from a Hodrick–Prescott trend, in the year that precedes a typical systemic banking crisis (histogram). The red line corresponds to the distribution in the full sample, which we use as benchmark. The figure shows that, before a systemic banking crisis both GDP and credit are above trend, with average deviations of

[5]The cross-correlations between credit and output over the sample also show significant differences between normal and crisis times. For example, during regular recessions the maximal correlation between (HP–filtered) credit and output is reached contemporaneously, with corr(credit$_t$,gdp$_t$)=0.38. We find similar results for periods outside recessions. In contrast, during financial recessions the maximal correlation is reached with one lag on credit (corr(credit$_{t-1}$,gdp$_t$)=0.24), suggesting that in those periods credit *leads* output.

1.8% and 3.8%, respectively. This suggests that crises break out at a particular point in the business cycle, typically in good times, in the midst of a credit boom. A general pattern extensively documented by Reinhart and Rogoff (2009, p. 157).

Figure 2: Distributions of GDP and credit gaps



3 The Model

We consider a closed economy populated with one representative risk averse household, one representative risk neutral competitive firm, and a mass one of heterogeneous, risk neutral, and competitive banks.

3.1 The Representative Firm

The representative firm lives for one period. It produces a homogeneous good that can be either consumed or invested by means of capital, k_t, and labor, h_t, according to a constant returns to scale technology represented by the production function[6] $F(k_t, h_t; z_t)$ where z_t is the level of total factor productivity (TFP), which is assumed to follow an AR(1) process of the form

$$\log z_t = \rho_z \log z_{t-1} + \varepsilon_t$$

where $|\rho_z| < 1$ and ε_t is an exogenous normally distributed TFP shock with zero mean and standard deviation σ_z that is realized at the beginning of period t. Variations in productivity are the only source of uncertainty and ε_t is realized at the beginning of period t, before the firm decides on its production plan. Capital, k_t, depreciates at rate $\delta \in (0, 1)$. The firm is born with no resources and must borrow k_t from the banks at a gross corporate loan rate R_t at the beginning of the period to be able to achieve production. The corporate loan is repaid at the end of the period. The firm also rents labor services from the household at rate w_t.

[6]The production function is increasing in both inputs, concave and satisfies Inada conditions.

The production plan is decided so as to maximize profits, which are given by

$$\pi_t = F(k_t, h_t; z_t) + (1-\delta)k_t - R_t k_t - w_t h_t. \tag{1}$$

3.2 The Representative Household

The infinitely lived representative household supplies inelastically one unit of labor per period[7] and has preferences over the flow of consumption, c_t, which are represented by the utility function:

$$\max_{\{a_{t+\tau+1}, c_{t+\tau}\}_{\tau=0}^{\infty}} \mathbb{E}_t \sum_{\tau=0}^{\infty} \beta^{\tau} u\left(c_{t+\tau}\right), \tag{2}$$

where $u\left(c_t\right)$ satisfies the usual regularity conditions,[8] $\beta \in (0,1)$ is the psychological discount factor, and $\mathbb{E}_t\left(\cdot\right)$ denotes the expectation operator which is taken over $\{\varepsilon_{t+\tau+1}\}_{\tau=0}^{+\infty}$. The household enters period t with assets, a_t, which she deposits in the banking sector and from which she receives a state contingent gross return r_t. There is no friction between the household and the banking sector and —since Modigliani and Miller's theorem applies— we cannot say anything as to whether a_t is made of bank deposits or bank equity. (This assumption will be relaxed in Section 6.2.[9]) The household earns unit wage w_t from supplying her labor and receives profits π_t from the firm. This income is then used to purchase the consumption good and transfer assets to the next period. Accordingly, the budget constraint is given by

$$c_t + a_{t+1} = r_t a_t + w_t + \pi_t. \tag{3}$$

The saving decision is determined by the standard arbitrage condition

$$u'\left(c_t\right) = \beta \mathbb{E}_t\left(u'\left(c_{t+1}\right) r_{t+1}\right). \tag{4}$$

Notice that, as will become clear shortly, there exists a positive wedge between banks' gross return on corporate loans (R_t) and the gross return on bank equity/assets (r_t). This wedge is due to inefficiencies in the banking sector.

3.3 The Banking Sector

The banking sector is at the core of the model and plays a non–trivial role because of two specific features. First, banks are heterogeneous with respect to their intermediation technology — some banks are more efficient than others, which potentially gives rise to an interbank

[7]This latter assumption is made for exposition purposes only and will be relaxed in the quantitative analysis (see Sections 4 and 5).

[8]In particular, we have $u'(c) > 0$, $u''(c) < 0$, $u(0) = \infty$ and $u'(\infty) = 0$

[9]To emphasize the fact that the household owns the banks and r_t is state contingent, we will in the meantime interchangeably refer to a_t as bank equity or deposits, and to r_t as the gross return on bank equity/deposits. Note also that we implicitly assume the existence of frictions between the household and the firm that prevent the household from financing the firm directly —this assumption is standard in the macro–literature. In the companion technical appendix of this paper, we also consider the case where the firm finances a fraction of its investment directly through the market.

market. It follows that banks have two types of activities. On the one hand they run traditional banking operations, which consist in collecting deposits/equity from households and lending the funds to the firm. In Shin (2008) and Shin and Shin (2011)'s language, these are "core" activities and, accordingly, bank deposits/equity are banks' core liabilities. On the other hand, banks also issue interbank claims ("non–core" assets/liabilities) so as to re–allocate assets toward the most efficient banks.[10] Second, the banking sector is subject to both asymmetric information and moral hazard problems, which impair the functioning of the interbank market.

3.3.1 Banks

There is a continuum of one–period,[11] risk–neutral, competitive banks that raise deposits/equity a_t from the household at the end of period $t-1$. At the time they raise deposits/equity, banks are identical and, in particular, have all the same size as they enter period t. At the beginning of period t, each bank draws a random bank–specific intermediation skill. Banks therefore become heterogeneous. Let p denote the bank with ability p, and assume that the ps are distributed over the closed interval $[0,1]$ with cumulative distribution $\mu(p)$, satisfying $\mu(0)=0$, $\mu(1)=1$, $\mu'(p)>0$. Bank p must pay an intermediation, dead–weight, cost $(1-p)R_t$ per unit of loan at the end of the period, so that its net return on each loan is pR_t. This cost reflects the bank's operational costs, for example, the cost of collecting corporate loans or monitoring the firm.[12] As an outside option, banks also have the possibility to invest assets in their own project. This project does not involve any intermediation cost but yields a lower, constant, and exogenous payoff γ per unit of good invested. Such an investment is inefficient, *i.e.* $\gamma < R_t$.[13] While there are several ways to interpret this outside option, we

[10] "The relevant distinction between core– and non–core liabilities can be seen as having to do with whether the claim is held by the ultimate domestic creditors (the domestic household sector). Repos and other claims held by banks on other banks can be regarded as non–core liabilities which are more volatile", Shin and Shin (2011, p. 3).

[11] Banks that operate in period t are born at the end of period $t-1$ and die at the end of period t. We will assume in a moment that banks are heterogeneous and that their types are private information. The assumption of one–period living banks is made to preserve this asymmetry of information over time. An alternative and equivalent approach would be to allow banks to live infinitely and, in order to rule out potential reputation effects, to assume that the types are randomly drawn afresh every period.

[12] This assumption is not crucial but convenient, because in this case bank heterogeneity is immaterial to the representative firm, which always pays its debt irrespective of the bank it borrows from. One could consider several alternative setups without loss of generality. For example, one could assume that there is a continuum of firms and that banks have different monitoring skills, which determine the probability that the projects of the firms they respectively lend to succeed. Typically, the firms borrowing from the skillful banks would then be able to repay their loan in full, while those borrowing from inefficient banks would default. We do not use this setup because we want to confine the inefficiencies within the banking sector and, by doing so, stay the closest possible to the textbook neoclassical model, where firms do not default.

[13] Indeed, if γ were strictly above R_t then banks would not finance the firm and, because of an Inada condition on the production function ($\lim_{k\to 0}\partial F_k(k,h;z)/\partial k = +\infty$), the marginal productivity of capital would be infinite, hence a contradiction. And the case $\gamma = R_t$ is ruled out by the existence of financial intermediation costs (see below). Notice that, since in the absence of the storage technology unused goods would depreciate at rate δ, the net return of storage is $\gamma-(1-\delta)$, which we assume is positive.

will refer to it as a storage technology.[14] An important aspect of this assumption is that the funds invested in this outside option cannot be used to finance the firm. This is key for the model to generate credit crunches.

Bank heterogeneity gives rise to an intra–periodic interbank market, where the least efficient banks lend to the most efficient ones at gross rate ρ_t, with $\gamma \leqslant \rho_t \leqslant R_t$.[15] Unlike corporate loans, interbank loans do not bear operational costs. Banks take the interbank rate ρ_t and the corporate loan rate R_t as given. Given these rates, bank p decides whether, and how much, it borrows or lends. Hereafter, we will refer to the banks that supply funds on the interbank market as "lenders" and to those that borrow as "borrowers". Let ϕ_t be the — endogenous and publicly observable — amount borrowed per unit of deposit/equity by a borrower p, with $\phi_t \geqslant 0$. In the rest of the paper, we will refer to ϕ_t as the "market/interbank funding ratio", defined as the ratio of market funding (non–core liabilities) to traditional funding (core liabilities). Then bank p's gross return on equity/assets is

$$r_t\,(p) \equiv \max\left\{pR_t\,(1+\phi_t) - \rho_t\phi_t, \rho_t\right\}. \tag{5}$$

It is equal to $pR_t\,(1+\phi_t) - \rho_t\phi_t$ when bank p borrows $\phi_t a_t$ from other banks at cost ρ_t and lends $(1+\phi_t)a_t$ to the firm for return pR_t. And it is equal to ρ_t when, instead, bank p does not do financial intermediation and lends to other banks. Bank p chooses to be a borrower when

$$pR_t\,(1+\phi_t) - \rho_t\phi_t \geqslant \rho_t \Longleftrightarrow p \geqslant \overline{p}_t \equiv \frac{\rho_t}{R_t}. \tag{PC}$$

Inequality (PC) is the participation constraint of bank p to the interbank market as borrower, rather than as lender, and pins down the type of the *marginal bank* $\overline{p}_t$ that is indifferent between the two options. Banks with $p < \overline{p}_t$ delegate financial intermediation to more efficient banks with $p \geqslant \overline{p}_t$. In a frictionless world, all banks with $p < 1$ would lend to the most efficient bank, so that $\overline{p}_t = 1$. This bank would have an infinite market funding ratio ($\phi_t \to +\infty$) and corner all assets; the economy would then reach the First Best allocation. The presence of two frictions on the interbank market — moral hazard and asymmetric information — prevents the economy from achieving First Best.

Moral Hazard: We assume that the proceeds of the storage technology are not traceable and cannot be seized by creditors. This implies that interbank loan contracts are not enforce-

[14] One could assume that the return of this activity varies over time. This would not affect our results as long as the return is not strictly positively correlated with the business cycle and the outside option can still be used as an insurance against adverse aggregate shocks. To fix ideas, one can think of this outside option as an intra–period home production activity or as a safe haven. One could also assume that the household has access to this storage technology. This would not affect our results either. Indeed, since γ is the return that even the worst bank ($p = 0$) can make in any state of the nature, the ex–post return on bank deposits/equity is always above that of storage (*i.e.* $\gamma < r_t$). Hence it would never be optimal for the household to use this technology.

[15] The interbank rate is the same for all borrowers, otherwise those that promise the lowest returns would not attract any lender. It has to be the case that $\rho_t \leqslant R_t$, otherwise no bank would be willing to borrow on the interbank market. Likewise, we have $\rho_t \geqslant \gamma$, otherwise no banks would be willing to lend.

able and that banks can walk away with the funds raised on the interbank market, without paying the interbank loans. Following the current practice (e.g. Hart, 1995, Burkart and Ellingsen, 2004), we refer to such opportunistic behavior as "cash diversion". When a bank diverts cash, the proceeds ultimately accrue to the shareholder — *i.e.* the household. The so–diverted cash is stored until the end of the period, and yields the return γ.[16] A bank that diverts $(1+\phi_t)\,a_t$ faces a diversion cost proportional to the size of the loan, and can only run away with $(1+\theta\phi_t)\,a_t$, for a net payoff of $\gamma\,(1+\theta\phi_t)\,a_t$. Parameter $\theta \in [0,1]$ reflects the cost of diversion, which is zero when $\theta = 1$ and maximal when $\theta = 0$. From a corporate finance literature viewpoint (*e.g.* Tirole, 2006), this is a standard moral hazard problem: *(i)* the gain from diversion increases with ϕ_t, *(ii)* the opportunity cost of diversion increases with bank efficiency p and *(iii)* with the corporate loan rate R_t. Features *(i)* and *(ii)* imply that efficient banks with "skin in the game" are less inclined to run away than highly leveraged and inefficient banks. Feature *(iii)* is similar to feature *(ii)*, but in the "time–series" (as opposed to "cross–sectional") dimension; it implies that banks are more inclined to run away when the return on corporate loans is low. This latter feature captures recent empirical evidence that banks tend to take more credit risk in such a situation (Maddaloni and Peydro, 2011).

Asymmetric Information: Lenders do not observe borrowers' skills —*i.e.* p is privately known— and therefore do not know borrowers' private incentives to divert cash. In this context, the loan contracts signed on the interbank market are the same for all banks. Neither ϕ_t nor ρ_t depends on p.[17]

By limiting the borrowing capacity of the most efficient bank ($p = 1$), moral hazard will give less efficient banks room to borrow; hence the positive wedge between R_t and r_t. Moral hazard is not enough to generate market freezes, though. For this we also need uncertainty about the quality –and therefore some adverse selection– of borrowers. Hence, both moral hazard and information asymmetry will be necessary to generate SBCs in the model.

Lenders want to deter borrowers from diverting. They can do so by limiting the quantity of funds that borrowers can borrow, so that even the most inefficient banks (*i.e.* those that

[16]Two comments are in order here. First, we will soon see that an incentive compatibility constraint will make sure that no bank diverts cash in equilibrium. Hence, cash diversion will be an out–of–equilibrium threat. Second, to be consistent, the return on cash diversion must not be strictly higher than γ. Otherwise, the diversion technology would dominate storage and would then be the relevant outside option for the banks.

[17]To see this, consider a menu of debt contracts $\{\rho_t(\widetilde{p}), \phi_t(\widetilde{p})\}_{\widetilde{p}\in[0,1]}$ intended for the borrowers of types $\widetilde{p}$s, and notice that lenders' arbitrage across these contracts requires that $\rho_t(\widetilde{p}) = \rho_t\ \forall \widetilde{p} \in [0,1]$. It is easy to see that such a menu of contracts cannot be revealing because any borrower p (*i.e.* with $pR_t > \rho_t$) claiming being of type $\widetilde{p}$ would make profit $r_t\,(\widetilde{p} \mid p) = pR_t + (pR_t - \rho_t)\,\phi_t\,(\widetilde{p})$ and pick the contract with the highest $\phi_t\,(\widetilde{p})$, independent of its type. It is equally easy to see that there is no revealing menu of equity contracts either. Indeed, consider a menu of equity contracts $\{\eta_t(\widetilde{p}), \phi_t(\widetilde{p})\}_{\widetilde{p}\in[0,1]}$, where $\eta_t(\widetilde{p})$ would be the share of retained earnings. Then the net profit of bank p would be $\eta_t(\widetilde{p})\,(1+\phi_t(\widetilde{p}))\,pR_t$ and, in equilibrium, this bank would pick the contract that yields the highest $\eta_t(\widetilde{p})\,(1+\phi_t(\widetilde{p}))$, independently of its own p.

should be lending) have no interest in demanding a loan and diverting it:

$$\gamma\left(1+\theta\phi_t\right) \leqslant \rho_t. \tag{IC}$$

This incentive compatibility constraint sets a limit to ϕ_t, which can therefore also be interpreted as lenders' funding tolerance, *i.e.* the limit market funding ratio above which a bank refuses to lend or, in Holmström and Tirole's language, the borrower's pledgeable income.[18] The program of bank $p \geqslant \overline{p}_t$ thus consists in maximizing its return on equity $r_t(p)$ (see (5) with respect to ϕ_t subject to constraint (IC). Proposition 1 below follows from the fact that, by construction —see (PC), the net return on interbank borrowing is strictly positive for borrowers.[19]

Proposition 1 (Optimal Interbank Funding Ratio) *The IC constraint binds at the optimum of the borrowing bank p, which thus exhausts its borrowing capacity:* $\phi_t = \frac{\rho_t-\gamma}{\gamma\theta}$.

The positive relationship between ϕ_t and ρ_t is a critical feature of the interbank funding ratio. When ρ_t increases, the net present value of corporate loans diminishes and only the most efficient banks remain on the demand side of the market. Since these banks have little private incentive to divert, lenders tolerate a higher interbank funding ratio (ϕ_t goes up). This is due to the negative (positive) externality that the marginal bank exerts on the other banks when she enters (leaves) the demand side of the market as, by having higher incentives to run away, she then raises (reduces) lenders' counterparty fears. In the limit case where $\rho_t = \gamma$, there is no demand for interbank loan because borrowers cannot commit to repay. The interbank funding ratio ϕ_t and the type of the marginal bank $\overline{p}_t$ fully describe banks' optimal decisions.

3.3.2 Interbank Market

The equilibrium of the interbank market is characterized by the gross return ρ_t that clears the market. We look for an equilibrium where $\rho_t > \gamma$ so that $\phi_t > 0$ and trade takes place. Since a mass $\mu\left(\overline{p}_t\right)$ of banks lend a_t, the aggregate supply of funds is equal to $\mu\left(\overline{p}_t\right) a_t$. Since a mass $1-\mu\left(\overline{p}_t\right)$ of banks borrow $\phi_t a_t$, aggregate demand is equal to $\left(1-\mu\left(\overline{p}_t\right)\right)\phi_t a_t$. The market clears when (using relations (PC) and Proposition 1):

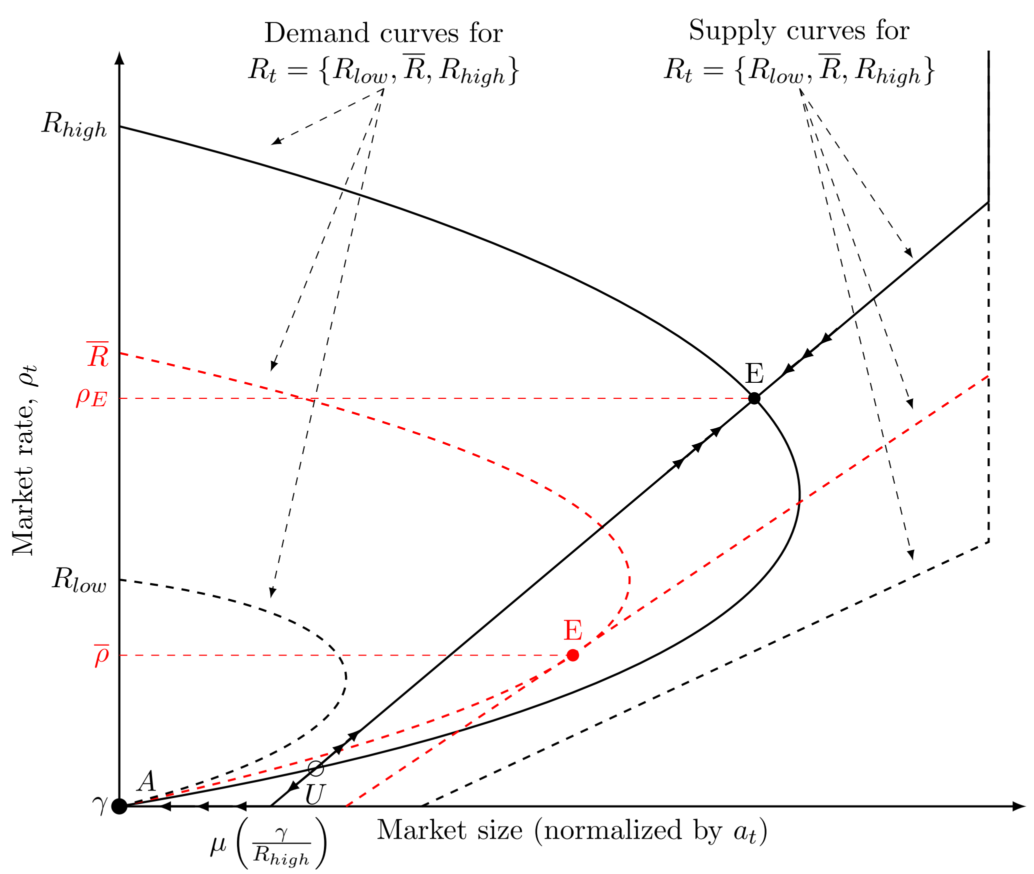
$$\underbrace{\mu\left(\frac{\rho_t}{R_t}\right)}_{\text{supply}} = \underbrace{\overbrace{\left(1-\mu\left(\frac{\rho_t}{R_t}\right)\right)}^{\text{extensive margin}}\;\overbrace{\frac{\rho_t-\gamma}{\gamma\theta}}^{\text{intensive margin}}}_{\text{demand}} \iff R_t = \Psi(\rho_t) \equiv \frac{\rho_t}{\mu^{-1}\left(\frac{\rho_t-\gamma}{\rho_t-\gamma(1-\theta)}\right)}. \tag{6}$$

[18] One could indeed recast the moral hazard problem into a setup *à la* Holmström and Tirole (1997), whereby borrowers may misuse the funds and enjoy private benefits at the expense of their creditors. *Stricto sensu*, the pledgeable income is the highest income that can be pledged without jeopardizing the borrower's incentives, *i.e.* $\rho_t(\rho_t-\gamma)a_t/\gamma\theta$.

[19] The proofs of all propositions are reported in Appendix A.

Aggregate supply increases monotonically with ρ_t, whereas aggregate demand is driven by two opposite forces. On the one hand, aggregate demand decreases with the interbank loan rate because fewer borrowers demand funds when the cost of funds increases; this is the "extensive margin" effect. On the other hand, a rise in ρ_t also exerts a positive effect on aggregate demand because each borrower is then able to borrow more; this is the "intensive margin" effect. At the aggregate level, this latter effect more than offsets the extensive margin effect when the marginal bank's externality affects a large mass of borrowers, *i.e.* when ρ_t is small enough. It follows that the aggregate demand curve binds backward, increasing with ρ_t for small values of ρ_t (see Figure 3). One can check that $\Psi(\rho_t)$ goes to infinity as ρ_t approaches γ, is greater than R_t when ρ_t approaches R_t, and reaches a minimum for some value $\rho_t = \overline{\rho} > \gamma$. Hence there exists a threshold $\overline{R} \equiv \Psi(\overline{\rho})$ for R_t below which there is no equilibrium with trade.[20] This threshold is the minimum corporate loan rate that is necessary for the banks to accept to lend to each other. Figure 3 illustrates this point and depicts the shifts in aggregate supply and demand as R_t falls from R_{high} (associated with equilibrium E) to R_{low}, with $R_{low} < \overline{R} < R_{high}$.

Figure 3: Interbank market clearing



[20] Equivalently, one can also write the market clearing condition in terms of $\overline{p}_t$ (since it is a multiple of ρ_t) and then obtain condition $\gamma\left(1-(\theta-1)\mu\left(\overline{p}_t\right)\right)/\overline{p}_t\left(1-\mu\left(\overline{p}_t\right)\right) = R_t$. It is easy to see that the left hand side expression is infinite for $\overline{p}_t = 0, 1$ and reaches a minimum $\overline{R}$ for some value $\overline{p}_t = \overline{\overline{p}} \in (0,1)$.

Following the fall in the corporate loan rate, the supply curve shifts to the right while the demand curve shifts to the left. Given the initial equilibrium rate $\rho_t = \rho_E$, demand falls below supply. Market clearing then requires that ρ_t go down, which results in more banks demanding funds (extensive margin). But since the banks that switch from the supply to the demand side are less efficient and have a relatively higher private incentive to divert cash, lenders require borrowers to deleverage. By construction, this intensive margin effect is the strongest when $R_t < \overline{R}$. It follows that, ultimately, aggregate demand decreases and excess supply goes further up. The de–leveraging process feeds itself and goes on until the market freezes, in point A, where $\rho_t = \gamma$.

In point E, where $R_t \geqslant \overline{R}$, borrowers have enough incentives to finance the firm and an interbank market equilibrium with trade exists; such a situation will be referred to as *normal times*. This equilibrium is *stable* in the sense that, in this point, net aggregate demand is a decreasing function of ρ_t and, following any small perturbation to ρ_t away from ρ_E, a standard Walrasian tatônnement process brings ρ_t back to ρ_E. For usual cumulative distributions, $\mu(p)$,[21] another interbank market equilibrium with trade is also possible, in point U. However, we rule it out because it is unstable.

In point A, where $R_t < \overline{R}$, things are different: autarky prevails. Demand and supply are both equal to zero, and the market clears because *(i)* borrowers have no pledgeable income ($\phi_t = 0$) and *(ii)* lenders are indifferent between interbank loans and storage. The marginal bank is then bank $\overline{p}_t = \gamma/R_t$, which is indifferent between financing the firm and using the storage technology. A mass $\mu(\gamma/R_t)$ of banks uses the storage technology, instead of lending to the firm. In the rest of the paper, we will interpret such a situation as a *systemic banking crisis*. This equilibrium is stable because net aggregate demand in this point decreases with ρ_t. Due to strategic complementarities between lenders (see Cooper and John, 1988), the autarkic equilibrium always exists, whatever the values of R_t. (Indeed, no bank has interest in making a loan if no one else does it.) Hence, it also always coexists with the equilibrium with trade whenever the latter exists. In order to rule out potential coordination failures we assume that banks always coordinate on the equilibrium with trade, which is Pareto–dominant, in this case.[22] That is, the interbank market freezes only when there exists no equilibrium with trade.

Based on relations (5) and (6), we can complete the description of the banking sector by

[21] Notably for the family distribution $\mu(p) = p^\lambda$ (with $\lambda \geqslant 0$) that we will be using later in the calibration.

[22] For a discussion on the selection of the Pareto–dominant equilibrium in games with multiple Pareto–rankable Nash equilibria, see Cooper et al. (1990).

deriving the sector's return on equity:

$$r_t \equiv \int_0^1 r_t(p)\,\mathrm{d}\mu(p) = \begin{cases} R_t \int_{\overline{p}_t}^1 p \frac{\mathrm{d}\mu(p)}{1-\mu(\overline{p}_t)} \text{ , if an equilibrium with trade exists} \\ \\ R_t \left(\frac{\gamma}{R_t}\mu\left(\frac{\gamma}{R_t}\right) + \int_{\frac{\gamma}{R_t}}^1 p\,\mathrm{d}\mu(p) \right) \text{ , otherwise.} \end{cases} \tag{7}$$

The interpretation of r_t is clear. When the equilibrium with trade exists, inefficient banks delegate financial intermediation to a mass $1-\mu(\overline{p}_t)$ of efficient banks, each of which therefore lending to the firm a multiple $1+\phi_t = 1/(1-\mu(\overline{p}_t))$ of their initial assets against net return pR_t. In autarky, in contrast, a mass $1-\mu(\gamma/R_t)$ of the banks make corporate loans, while the remainder use the storage technology. The banking sector is fully efficient when $\gamma \to 0$, *i.e.* when interbank loan contracts are fully enforceable, as in this case $\overline{R} \to 0$ and $\overline{p}_t \to 1$ (the interbank market always exists and only the best bank does the intermediation), and $\phi_t \to +\infty$ (the best bank is infinitely leveraged). The same is true when $\lim_{p\nearrow 1}\mu(p) = 0$, since in this case there is a mass one of banks with $p = 1$ and banks are homogeneous and all efficient.

3.3.3 Aggregate Supply of Corporate Loans

In normal times banks reallocate their assets through the interbank market, and all assets a_t are ultimately channeled to the firm. In crisis times, in contrast, the interbank market freezes and only the banks with $p \geqslant \gamma/R_t$ lend to the firm. As a consequence, the banking sector only supplies $(1-\mu(\gamma/R_t))\,a_t$ as corporate loans. Denoting by k_t^s banks' aggregate supply of corporate loans, one thus gets:

$$k_t^s = \begin{cases} a_t \text{ , if an equilibrium with trade exists} \\ \\ \left(1-\mu\left(\frac{\gamma}{R_t}\right)\right) a_t \text{ , otherwise} \end{cases} . \tag{8}$$

3.4 Recursive Decentralized General Equilibrium

A general equilibrium of the economy is defined as follows.

Definition 1 (Recursive decentralized general equilibrium) *A decentralized recursive general equilibrium is a sequence of prices $\mathscr{P}_t \equiv \{R_{t+i}, r_{t+i}, \rho_{t+i}, w_{t+i}\}_{i=0}^{\infty}$ and a sequence of quantities $\mathscr{Q}_t \equiv \{c_{t+i}, y_{t+i}, k_{t+i}, h_{t+i}, a_{t+i}\}_{i=0}^{\infty}$ such that for a given sequence of prices, $\mathscr{P}_t$, the sequence of quantities, $\mathscr{Q}_t$, solves the optimization problems of the agents, and for a sequence of quantities, $\mathscr{Q}_t$, the sequence of prices, $\mathscr{P}_t$, clears the markets.*

In equilibrium, the household supplies one unit of labor, implying that the production level is given by $f(k_t; z_t) \equiv F(k_t, 1; z_t)$ and the marginal efficiency of capital is $f_k(k_t; z_t) \equiv \frac{\partial F(k_t,1;z_t)}{\partial k_t}$.

The market clearing condition on the corporate loan market thus takes the form

$$f_k^{-1}(R_t + \delta - 1; z_t) = \begin{cases} a_t \text{ , if an equilibrium with trade exists} & (a) \\ a_t - \mu\left(\frac{\gamma}{R_t}\right) a_t \text{ , otherwise.} & (b) \end{cases} \tag{9}$$

Relation (9) yields the equilibrium R_t as a function of the two state variables of the model, a_t and z_t. It also points to the two–way relationship that exists between the interbank loan market and the retail corporate loan market. We indeed showed that the way the interbank operates depends on whether or not $R_t \geqslant \overline{R}$. Likewise, whether or not the interbank market operates has an impact on the supply of corporate loans and, therefore, on R_t. To solve for the general equilibrium we need to take into account these feedback effects. We proceed in two steps. First, we solve (9a) for R_t under the conjecture that the interbank market equilibrium with trade exists, and then check a posteriori whether indeed $R_t \geqslant \overline{R}$. In the negative, the interbank market equilibrium with trade cannot emerge, and the interbank market freezes. In this case the equilibrium corporate loan rate is the R_t that solves (9b). Proposition 2 follows.

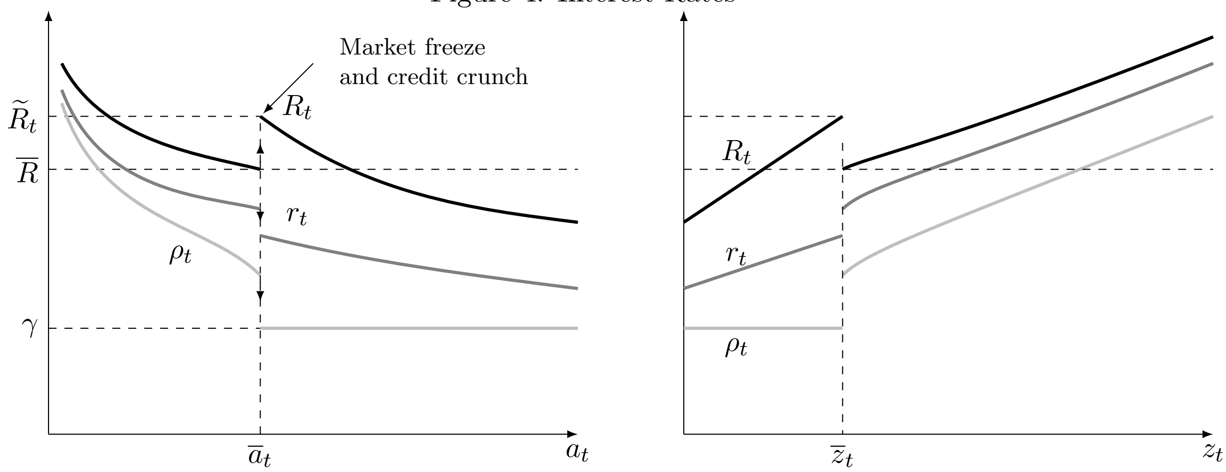
Proposition 2 (Interbank loan market freeze) *The interbank loan market is at work if and only if* $a_t \leqslant \overline{a}_t \equiv f_k^{-1}(\overline{R} + \delta - 1; z_t)$*, and freezes otherwise.*

The threshold $\overline{a}_t$ is the maximum quantity of assets that the banking sector can reallocate efficiently. Above this threshold counterparty fears on the interbank market are so widespread that mistrust prevails and the interbank market freezes. In the rest of the paper we will refer to $\overline{a}_t$ as the absorption capacity of the banking sector. Importantly, Proposition 2 suggests that the ability of the banking sector to re–allocate assets internally ultimately depends on the level of productivity in the real sector, z_t. The more productive the real sector, the more efficient the banking sector ($\partial \overline{a}_t / \partial z_t > 0$). The intuition and mechanics are clear. An increase in total factor productivity raises the demand for capital and the equilibrium corporate loan rate. By raising banks' opportunity cost of storage and cash diversion, the increase in R_t also reduces uncertainty about counterparties' quality, making it less likely for the interbank loan market to freeze. Given a level of assets a_t, there therefore exists a productivity threshold $\overline{z}_t$ below which the interbank market freezes (with $\overline{z}_t \equiv f_k^{-1}(\overline{R} + \delta - 1; a_t)$ and $\partial \overline{z}_t / \partial a_t < 0$). Overall, our model captures the notion that banks' core liabilities (equity/deposits a_t), which are predetermined, are a stable source of funding whereas non–core liabilities are unstable funding because they are subject to market runs. Proposition 3 below shows how disruptions in the wholesale financial market spill over the retail loan market and trigger a credit crunch.

Proposition 3 (Credit crunch) *An interbank market freeze is accompanied with a sudden fall in the supply of corporate loans* k_t^s *(i.e. given* z_t*,* $\lim_{a_t \searrow \overline{a}_t} k_t^s < \lim_{a_t \nearrow \overline{a}_t} k_t^s$*), as well* *as by a sudden increase in the interest rate spread R_t/r_t (i.e. given z_t, $\lim_{a_t \searrow \overline{a}_t} R_t/r_t > \lim_{a_t \nearrow \overline{a}_t} R_t/r_t$). We will refer to such a situation as a credit crunch.*

Panel (a) of Figure 4 illustrates Proposition 3. It depicts the equilibrium rates, R_t, r_t, and ρ_t as functions of a_t for a given level of productivity z_t. The corporate loan rate monotonically decreases with bank assets almost everywhere, but there is a break for $a_t = \overline{a}_t$, when the level of assets reaches the banking sector's absorption capacity. Above this threshold, a credit crunch occurs and the corporate loan rate suddenly jumps to $\widetilde{R}_t \equiv \lim_{a_t \searrow \overline{a}_t} R_t$, with $\widetilde{R}_t > \overline{R}$. Notice that, from a *partial* equilibrium perspective, $\widetilde{R}_t$ is high enough to restore banks' incentives and reignite the interbank market. But this situation is not sustainable as a *rational expectation general* equilibrium, since by issuing interbank claims banks would be able to raise their supply of corporate loans and, ultimately, R_t would go down below $\overline{R}$. It follows that the autarkic equilibrium is the only interbank market equilibrium that is consistent with the general equilibrium when $a_t > \overline{a}_t$.[23] Panel (b) of Figure 4 depicts the equilibrium rates as functions of z_t for a given level of assets a_t and mirrors Panel (a). It shows that a crisis breaks out as soon as z_t falls below $\overline{z}_t$. Altogether, these two figures suggest that a SBC may result either from the *endogenous* over–accumulation of assets by the household (Panel (a)), or from an *exogenous* adverse productivity shock (Panel (b)). This variety of banking crises is an important and novel feature of our model, which we will discuss in detail in Sections 4.2 and 5.

Figure 4: Interest Rates



(a) Interest Rates vs Assets (b) Interest Rates vs Technology Shock

Two regimes prevail in the economy. The first regime corresponds to normal times, where the interbank market functions well and all the assets available in the economy are used to

[23]In other words, condition $R_t \geqslant \overline{R}$ is necessary but not sufficient to rule out interbank market freezes. In a world with monopolistic banks (e.g. *à la* Stiglitz and Weiss, 1981), banks may have an interest in rationing their supply of corporate loans so as to maintain R_t above $\overline{R}$ and keep access to the interbank market. They cannot do this here because they are atomistic and price takers.

finance the firm. The other regime corresponds to crisis times, where the interbank market shuts down, preventing the efficient re-allocation of assets. In this case, bank inefficiencies materialize themselves as a widening of the interest rate spread, which is due to the simultaneous increase in the corporate loan rate and the fall in the return on bank asset/equity. Which regime prevails depends on the size of the banks relative to their absorption capacity and on whether bank lending is "excessive". In other words, credit booms may be bad. To tell apart the bad from the good credit booms, we define the probability of a crisis at a n-period horizon as the joint probability that the banking sector's total assets exceed its absorption capacity in $t+n$ (*i.e.* that $a_{t+n} > \overline{a}_{t+n}$) and not before (*i.e.* $a_{t+i} \leqslant \overline{a}_{t+i}$ for $i = 1, ..., n-1$).

Definition 2 (Probability of a crisis at a $n-$ period horizon) *Given the data generating process of productivity, the state of the nature at the end of period t (a_{t+1}, z_t) and the optimal asset accumulation rule $a_{t+i+1}(a_{t+i}, z_{t+i})$, the probability that a systemic banking crisis next breaks out in period $t+n$ is $\mathbb{P}(a_{t+1} \leqslant \overline{a}_{t+1}, \dots, a_{t+n-1} \leqslant \overline{a}_{t+n-1}, a_{t+n} > \overline{a}_{t+n}) = \int_{\overline{\varepsilon}_{t+1}}^{+\infty} \dots \int_{\overline{\varepsilon}_{t+n-1}}^{+\infty} \int_{-\infty}^{\overline{\varepsilon}_{t+n}} \mathrm{d}G(\varepsilon_{t+1}, \dots, \varepsilon_{t+n})$, where $G(\cdot)$ denotes the cumulative of the n-variate Gaussian distribution, $n > 1$, and $\overline{\varepsilon}_{t+i} \equiv \ln \overline{z}_{t+i} - \rho_z \ln z_{t+i-1}$, where $\overline{z}_{t+i}$ is the threshold of productivity in period $t+i$ below which, given the level of financial assets a_{t+i}, a crisis breaks out — i.e. $\overline{z}_{t+i} \equiv f_k^{-1}(\overline{R} + \delta - 1; a_{t+i})$.*

The crisis probability provides an early warning signal of banking crises. Of course, other variables in the model could be used to construct more standard indicators, like the credit to output ratio or the growth rate of credit. But these statistics do not contain as much information about future crises as the crisis probability does, because the latter is fully consistent with general equilibrium effects, agents' rational expectations, and perceived risks. For instance, the ex–ante anticipation of a market freeze leads the household to accumulate assets faster so as to smooth consumption should the market indeed freeze. By doing so, however, the household feeds a supply–driven credit boom, making the crisis more likely ex post. Hence the high crisis probability ex ante.

4 Calibration and Solution of the Model

We assess quantitatively the ability of the model to account for the dynamics of the economy before, during, and after a banking crisis. To do so, we extend the model to the presence of endogenous labor supply decisions.[24] The technology is assumed to be represented by the standard constant–returns to scale Cobb–Douglas function, $F(k_t, h_t; z_t) \equiv z_t k_t^{\alpha} h_t^{1-\alpha}$ with $\alpha \in (0, 1)$. The household is assumed to be endowed with preferences over consumption and

[24]The equations characterizing the general equilibrium in this case are reported in Appendix D.1.

leisure that are represented by the following Greenwood et al. (1988) utility function

$$\mathbb{E}_t \sum_{\tau=0}^{\infty} \beta^{\tau} \frac{1}{1-\sigma} \left(c_{t+\tau} - \vartheta \frac{h_{t+\tau}^{1+\upsilon}}{1+\upsilon} \right)^{1-\sigma}.$$

It is well known that these preferences wipe out wealth effects in labor supply decisions, and we chose this specification for practical reasons. By yielding a closed form solution for the absorption capacity of banks

$$\overline{a}_t = \Gamma z_t^{\frac{1+\upsilon}{\upsilon(1-\alpha)}}, \text{ with } \Gamma \equiv \left(\frac{1-\alpha}{\vartheta}\right)^{\frac{1}{\upsilon}} \left(\frac{\alpha}{\overline{R}+\delta-1}\right)^{\frac{\upsilon+\alpha}{\upsilon(1-\alpha)}},$$

that remains independent of the consumption level, it indeed preserves the sequential character of the resolution of the model. (As we will show later, this assumption plays against our model as the absence of a wealth effect on labor will be one reason why our model overstates the amplitude of banking crises compared with the data.) The absorption capacity now depends in a fundamental way on υ, the (inverse of) labor supply elasticity. The larger labor supply elasticity, the more elastic to productivity shocks the absorption capacity of banks. The reason is that, in this case, positive productivity shocks entail a larger increase in hours worked, hence a larger increase in the marginal productivity of capital, in the corporate loan rate, and ultimately in the absorption capacity.

4.1 Calibration

The model is calibrated on a yearly basis, in line with Jordà et al. (2011a)'s database, which we used to document the stylized facts on banking crises in Section 2. We use a rather conventional calibration of the model (see Table 2). The discount factor, β, is set such that the household discounts the future at a 3% rate per annum. The inverse labor supply elasticity is set to $\upsilon = 1/3$ which lies within the range of values that are commonly used in the macro literature. The labor disutility parameter, ϑ, is set such that the household would supply one unit of labor in the average steady state. The curvature parameter σ is set to 4.5, which lies within the range of estimated values for this parameter.[25] However, given the importance of this parameter for the dynamics of precautionary savings, we will assess the sensitivity of our results to changes in this parameter. The capital elasticity in the production function is set to $\alpha = 0.3$ and capital is assumed to depreciate at a 10% rate per annum ($\delta = 0.1$). The process for the technology shock is estimated using the annual total factor productivity series produced by Fernald (2009) for the US over the post–WWII period. We obtain a persistence parameter $\rho_z = 0.9$ and a standard deviation $\sigma_z = 1.81\%$.

The remaining parameters pertain to the banking sector and include the return on storage γ, the cost of diversion θ, and the distribution of banks $\mu(\cdot)$. For tractability reasons we assume

[25] Mehra and Prescott (1985), for example, impose an upper bound of 10.

that $\mu(p) = p^\lambda$, with $\lambda \in \mathbb{R}^+$. The parameters of the banking sector are calibrated jointly so that in the simulations the model generates *(i)* a systemic banking crisis every forty years on average, *i.e.* with probability around 2.5% (see Table 1), *(ii)* an average interest rate spread equal to 1.71%, and *(iii)* an average corporate loan rate of 4.35%.[26] These latter two figures correspond to the lending rate on mid–size business loans for the US between 1990 and 2011, as reported in the US Federal Reserve Bank's Survey of Terms of Business Lending —see Appendix C.2. We obtain $\gamma = 0.94$, $\lambda = 24$, and $\theta = 0.1$. Based on this calibration, the model generates an average interbank loan rate of 0.86% and an implied threshold for the corporate loan rate of 2.43% (*i.e.*, $\overline{R} = 1.0243$).

Table 2: Calibration

Parameter		Values
Discount factor	β	1/1.03
Risk aversion	σ	4.500
Frish elasticity	υ	1/3
Labor disutility	ϑ	0.944
Capital elasticity	α	0.300
Capital depreciation rate	δ	0.100
Standard dev. productivity shock	σ_z	0.018
Persistence of productivity shock	ρ_z	0.900
Bank distribution; $\mu(p) = p^\lambda$	λ	24
Diversion cost	θ	0.1
Storage technology	γ	0.936

4.2 Optimal Asset Accumulation Rule

The model is solved by a collocation method.[27] We first discretize the technology shock using the method developed by Tauchen and Hussey (1991), with 31 nodes for the Gauss–Hermite approximation. The decision rule for a_{t+1} is approximated by a function of Chebychev polynomials of the form

$$a_{t+1}(a_t; z_i) = \exp\left(\sum_{j=0}^{q} \xi_j^{\mathrm{N}}(z_i) T_j(\varphi(a_t))\right) \mathbb{I}_{a_t \leqslant \overline{a}_t} + \exp\left(\sum_{j=0}^{q} \xi_j^{\mathrm{C}}(z_i) T_j(\varphi(a_t))\right) \mathbb{I}_{a_t > \overline{a}_t},$$

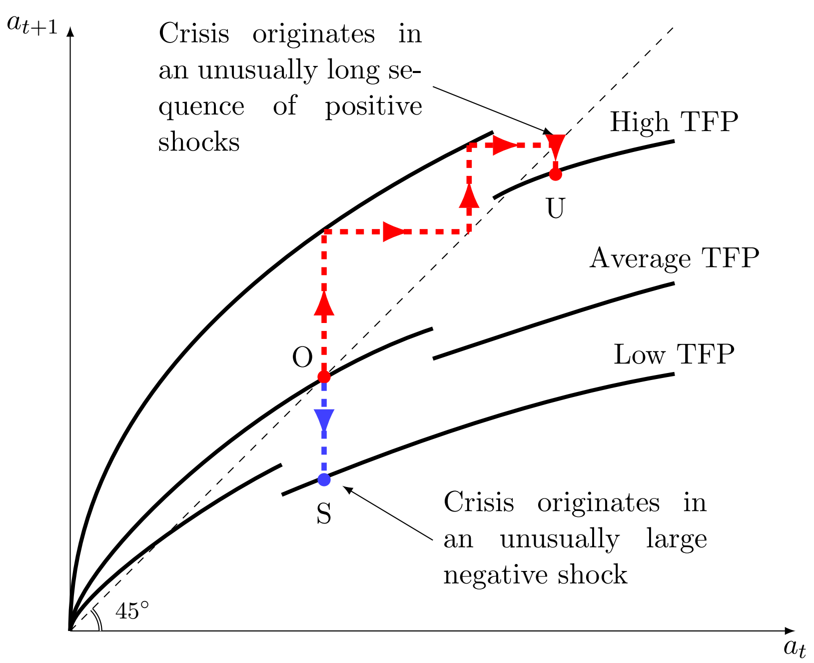
where z_i denotes a particular level of the total factor productivity in the grid. $T_j(\cdot)$ is a Chebychev polynomial of order j, $\xi_j^{\mathrm{N}}(z_i)$ (resp. $\xi_j^{\mathrm{C}}(z_i)$) is the coefficient associated to this polynomial when the economy is in normal times (resp. in a systemic banking crisis) for the value of productivity z_i. We use 15th order Chebychev polynomials (q=15). Finally,

[26]Because precautionary savings play an important role in our model there is a significant gap between the deterministic and the stochastic steady states. It is therefore more accurate to calibrate the model based the moments calculated over the simulations of the model, rather than on the basis of the deterministic steady state.

[27]The companion technical appendix provides greater details on this solution method.

$\varphi(a_t)$ is a function that maps the level of assets into the interval (-1,1),[28] and $\mathbb{I}_{a_t \leqslant \bar{a}_t}$ (resp. $\mathbb{I}_{a_t > \bar{a}_t}$) is an indicator function that takes value one in normal (resp. crisis) times and zero otherwise. We allow for a discontinuity in the rule at the points when total assets reach the banking sector's absorption capacity —that is, when $a_t = \bar{a}_t$— because in those points the economy switches regime. (That $\bar{a}_t$ is a known function of state variable z_t greatly simplifies the solution algorithm.) The optimal decision rule $a_{t+1}(a_t; z_t)$ is given by the fixed point solution to the Euler equation (4). The parameters of the approximated decision rule are set such that the Euler residuals are zero at the collocation nodes. We checked the accuracy of our solution using the criteria proposed by Judd (1992).

Figure 5: Optimal Decision Rules



Note: Dashed line: 45^o line where $a_{t+1} = a_t$.

Figure 5 illustrates the optimal asset accumulation decision rules for the lowest, the average, and the highest productivity levels in our grid.[29] As in the standard, frictionless neoclassical model, the household smooths her consumption over time by accumulating relatively more (less) assets when productivity is above (below) average. But the pace of accumulation is faster than in the frictionless model because of the possibility of crises, which lead the household to also save for precautionary motives. We will discuss this point in the next section.

[28]More precisely, we use $\varphi(a_t) = 2\frac{\log(a_t)-\log(y)}{\log(x)-\log(y)} - 1$, where $x \in \{a_{sup}, \bar{a}_i\}$ and $y \in \{a_{min}, \bar{a}_i\}$. a_{min} and a_{sup} denote the bounds of the interval values for a_t (we use $a_{min} = 0.5$ and $a_{sup} = 20$), z and y denote respectively the upper and the lower bound values we use in normal times and during a systemic banking crisis.

[29]For illustrative and exposition purposes, we report stylized representations of the decision rules. The solution decision rules for the benchmark calibration are reported in the companion technical appendix.

The optimal decision rules provide a first insight into the dynamics of the model. Starting from the average steady state O, the economy may face a systemic banking crisis for two different reasons. It may first experience a large negative technology shock, which brings the economy down to S. Because this shock instantaneously reduces banks' absorption capacity below the current level of assets, the crisis breaks out on impact. Such a crisis is purely driven by the bad realization of the exogenous shock and the amplification mechanism that the shock triggers. This is the channel usually depicted in the existing literature. In our model, a crisis may occur for another reason. As the economy experiences an unusually long sequence of TFP levels above average, the household has time to accumulate a large stock of assets, which may outgrows banks' absorption capacity, as in point U. In this case the crisis results from the optimal response of the household to positive, as opposed to negative, events. Next section describes such credit–boom led crises in more details.

5 Quantitative Analysis

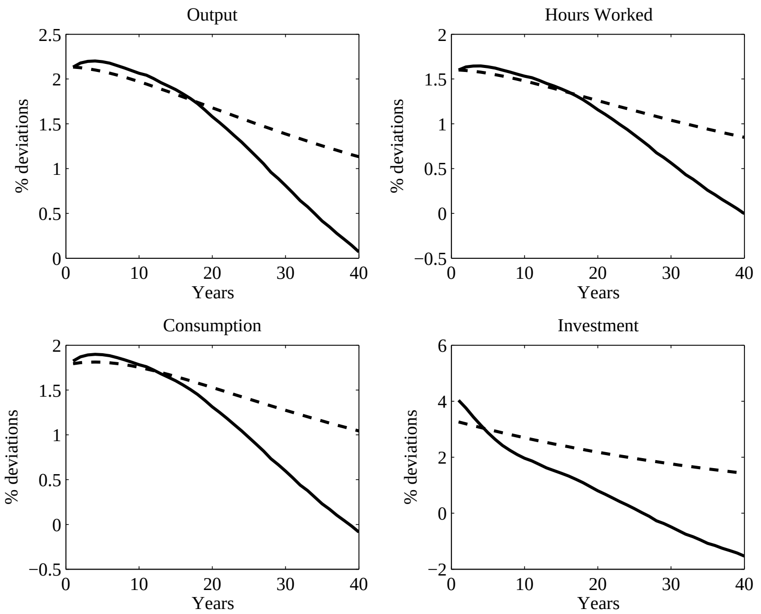
5.1 The Economy in Normal Times

We start by analyzing the dynamics of the economy in response to a positive one standard deviation productivity shock about the average steady state. Figures 6–7 report the dynamics of our model (plain line) and those as obtained from the basic frictionless neoclassical model (dashed line), which corresponds to our model when $\gamma = 0$, $\theta = 0$, or $\lambda = +\infty$.[30]

Figure 6 reveals that the model behaves very similarly to the standard frictionless model in the face of a small positive shock. In both models, output, consumption, investment and hours worked increase on impact. After one period, the dynamics of the two models depart. The dynamics generated by our model is hump–shaped, whereas the frictionless economy goes back monotonically to the steady state. The hump–shaped pattern comes from a financial accelerator mechanism similar in spirit to that in Bernanke et al. (1999) that amplifies the effects of the shock. Indeed, the corporate loan rate rises on impact (see Figure 7), which mitigates counterparty fears on the interbank loan market and relaxes banks' borrowing constraints. As every borrower raises its market funding, the aggregate demand for interbank loans increases and so does the interbank loan rate. Inefficient banks then switch from the demand to the supply side of the market. This works to raise borrowers' overall quality and reduces the moral hazard problem further. Unlike Bernanke et al. (1999), however, the financial accelerator mechanism comes from frictions on the wholesale loan market —not from frictions on the retail loan market. The response of financial variables in Figure 7 makes

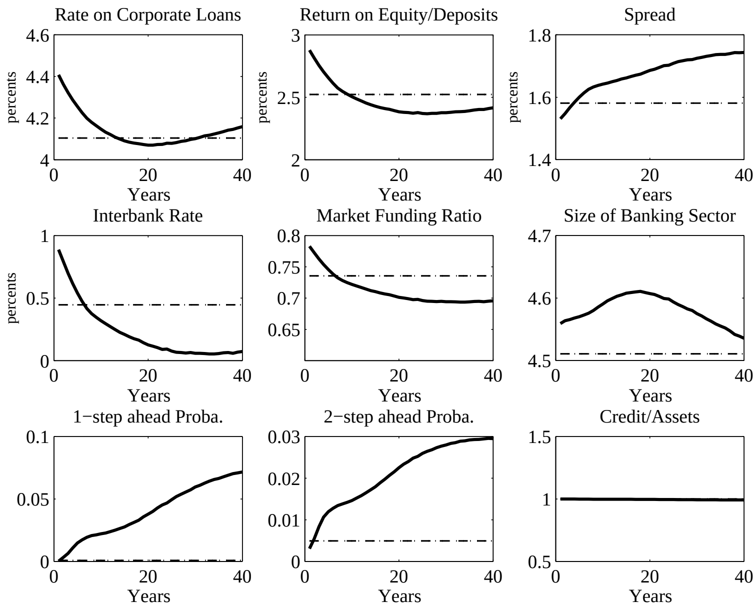
[30]We report the average of the distribution of the dynamic paths in the economy, as obtained from 100,000 simulations of the model. Notice that, despite the existence of a storage technology, our frictionless model is very similar to –and in effect indistinguishable from– the basic neoclassical growth model. The reason is that for this technology to be operated, one would need an unrealistic, zero–probability, negative TFP shock to occur.

Figure 6: Impulse Response to a One Standard Deviation Technology Shock (I)



Note: Plain line : Average impulse response function across 100,000 simulations in the model, Dashed line: Average Impulse Response across simulations in the frictionless economy.

Figure 7: Impulse response to a 1 Standard Deviation Technology Shock (II)



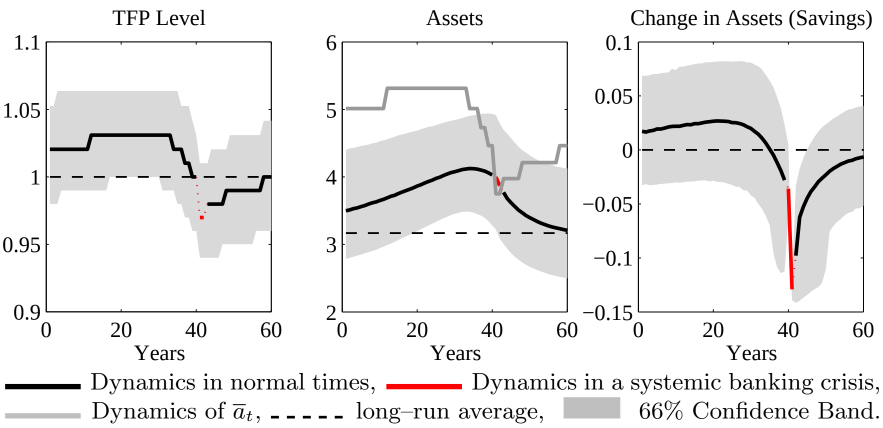
Note: Plain line : Average path across 100,000 simulations in the model, Dashed line: Steady state value.

it clear that the economy does not experience any systemic banking crisis. Even though the corporate loan rate eventually falls below its steady state value as the household accumulates assets, at no point in the dynamics does it fall below $\overline{R}$ (*i.e.* 2.43%). In effect, the positive technology shock does not last long enough to have the household accumulate assets beyond banks' absorption capacity. We obtain similar (mirror) results after a negative one standard deviation productivity shock. Most of the time, the model behaves like a standard financial accelerator model; crises are indeed rare events that occur under specific conditions, as we show in the next section.

5.2 Typical Path to Crisis

The aim of this section is to describe the typical conditions under which systemic banking crises occur. As we already pointed out (see Section 4.2), banking crises may break out in bad as well as in good times in the model. It is therefore not clear which type of shocks (negative/positive, large/small, short/long lived) are the most conducive to crises. Starting from the average steady state (*i.e.* $z_t = 1$), we simulate the model over 500,000 periods, identify the years when a crisis breaks out, and compute the median underlying technological path in the 40 (resp. 20) years that precede (resp. follow) a crisis. This path corresponds to the typical sequence of technology shocks leading to a crisis. We then feed the model with this sequence of shocks. The left panel of Figure 8 reports the typical path for the technology

Figure 8: Typical path (I)



shock, along with its 66% confidence interval. The red part of the path corresponds to crisis periods, the black one is associated with normal times. One striking result that emerges from this experiment is that the typical banking crisis is preceded by a long period during which total factor productivity is above its mean. In some 20% of the cases, crises even occur at a time when productivity is still above mean. This reveals one important and interesting aspect of the model: the seeds of the crisis lie in productivity being above average for an unusually long time. The reason is that a long period of high productivity gives the household enough time to accumulate assets beyond the banking sector's absorption capacity (see middle panel in Figure 8). This is a phase where financial imbalances build up, as reflected in the increase in the one–step ahead probability from an initial 0 to 0.25 few years before the banking crisis in Figure 9.[31]

In our model, the typical crisis breaks out in the middle of a credit boom. At the beginning both productivity and the corporate loan rate are above their average steady state, and the credit boom is demand–driven. As productivity goes back to average, the firm reduces its demand for labor, capital, and credit. Following the decrease in her labor income, the household reduces her savings and even dissaves (see right panel in Figure 8). But this reduces only marginally the household's total stock of assets and the overall supply of credit by banks, which remains more than 30% above steady state level. At this point in time the credit boom turns supply–driven and the corporate loan rate goes down below its average steady state (see Figure 9).[32] The fall in the corporate loan rate is followed by a fall in the interbank loan rate, which gives inefficient banks incentives to demand interbank loans. This has two important consequences. First, since more banks finance themselves on the market, the aggregate balance sheet of the banking sector increases in size.[33] Second, the fact that more inefficient banks raise funds erodes trust between banks. It follows that the banking sector gets bigger precisely at the time when its absorption capacity diminishes, making it less resilient to adverse shocks. In this context, the rather mild 3% drop of productivity below its average in year 40 is enough to bring the whole sector down. The crisis manifests itself as a sudden 10% fall in the credit to assets ratio, k_t/a_t, and a 30% drop in the size of the banking sector, $a_t + (1 - \mu(\overline{p}_t))\,\phi_t a_t$. As banks cut their supply of corporate loans, the spread rises from 2% to almost 4%. Interestingly, the rise in the spread is due to the return on equity/deposits falling by more than the corporate loan rate during the crisis, a phenomenon for example observed in 2007–8 in the US.[34] The dynamics of the corporate loan rate results from two opposite effects. On the one hand, the crisis brings about a credit crunch that implies, all other things equal, a rise in the corporate loan rate (see Figure 4). On the other hand, the fall in productivity also implies, in the first place, a fall in the firm's demand for corporate loan, and therefore a decrease in the corporate loan rate. In the typical crisis, this latter effect dominates. Ultimately, the corporate loan rate decreases, but by

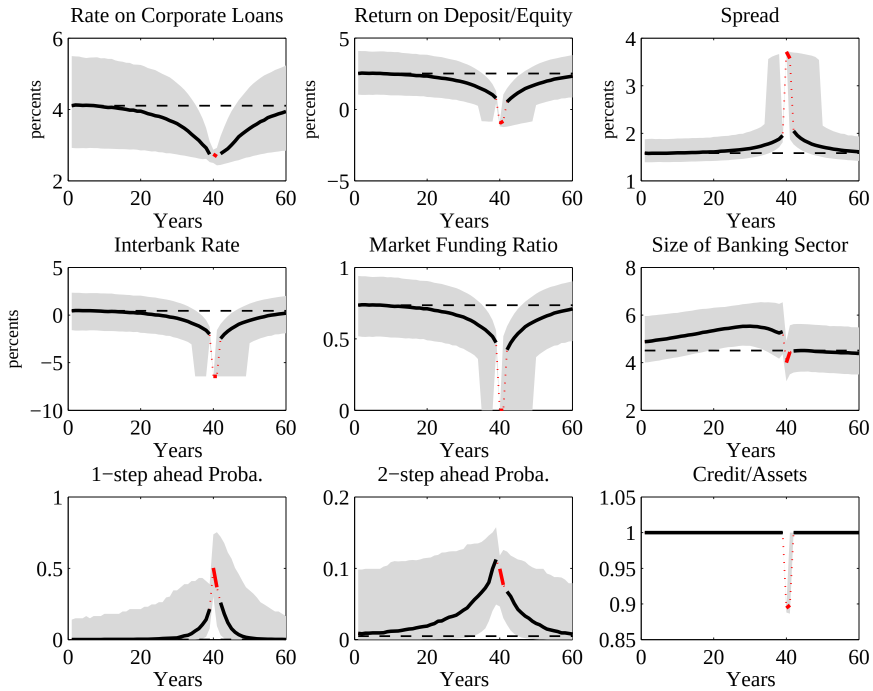
[31] Likewise, the two–step ahead probability goes up to 0.15. These probabilities are constructed following Definition 2. Hence, the two–step ahead probability in period t is the probability that a banking crisis will break out in period $t+2$, conditional on the event that there is no crisis in period $t+1$.

[32] To the best of our knowledge —because of lack of data— there exists no clear historical pattern of the dynamics of interest rates before SBCs. However, the low level of real US interest rates has been shown to be one important factor of the recent crisis (Taylor, 2009). The latter was indeed preceded by particularly low real fed fund rates, with an average of 0.84% over 2000-2007 despite the monetary policy tightening in 2005-2006, against 3.24% over 1986-2000; see Figure C.4 in Appendix C.2.

[33] On the brink of the typical crisis, the total assets of the banking sector are about 30% above steady state. In the run up to the crisis, borrowing banks' market funding ratio (ϕ_t) goes down so as to maintain private incentives, which is consistent with the pre-2007 crisis developments for the US. But since more banks become leveraged, the leverage of the banking sector as a whole, $(1 - \mu(\overline{p}_t))\,\phi_t$, increases.

[34] See Figure C.4 in Appendix C.2.

Figure 9: Typical path (II)



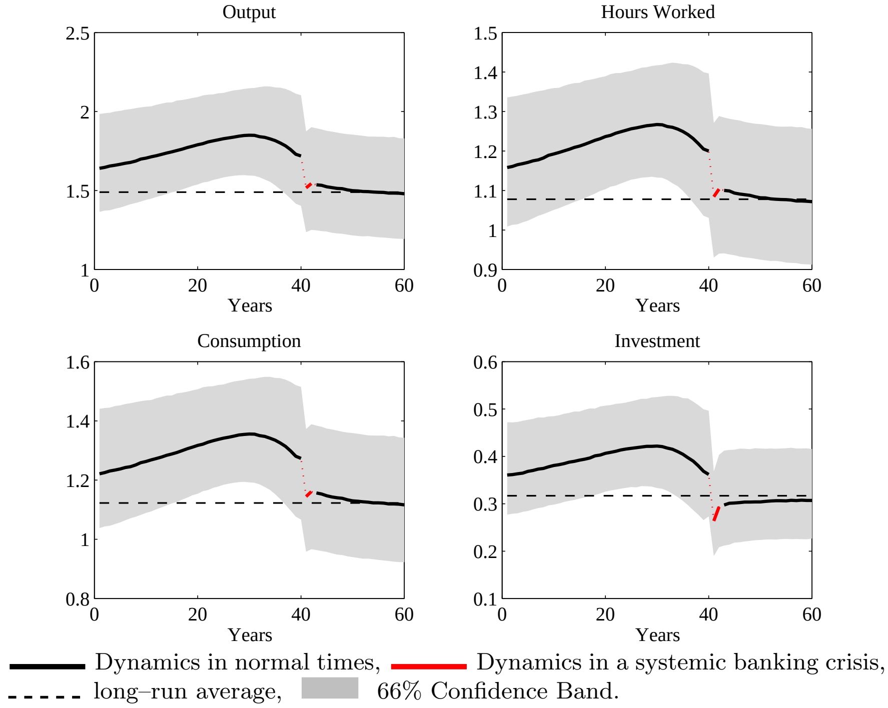
Dynamics in normal times, Dynamics in a systemic banking crisis, long–run average, 66% Confidence Band.

much less than it would otherwise have in normal times following a similar —3%— drop in productivity. After the crisis, all financial variables return back to their steady state levels. Finally, Figure 10 illustrates the evolution of macroeconomic variables. In the run up to the crisis, the positive wealth effect associated with technological gains leads the household to both consume and invest more. The household also increases her labor supply in order to take advantage of the high real wage, so that the level of output is 15% above its steady state in the year before the crisis. These unusually good times end abruptly with the credit crunch, which triggers a sizable drop in investment, labor, consumption, and output. In the year of the crisis, output falls by 15%. Part of this loss —around 3%— is due to the fall in productivity that takes place as the crisis hits, and would have taken place even in the absence of the crisis as the result of the standard mean–reverting dynamics of productivity; the remaining 12% of the fall in output is attributable to the banking crisis *per se*. Notice that output falls by more than in the data (6.74% —see Table 1), suggesting that financial frictions in the model are relatively severe.[35] However, the performance of the model can be improved on that front *(i)* by allowing the firm to issue corporate bonds or equity directly to the household so as to reduce the prevalence of the banking sector in firm financing, or *(ii)* by

[35]For instance, the assumption that banks live for only one period makes information asymmetries particularly prevalent. One way to mitigate the information problem would be to have banks living several periods with persistent skills. (An extreme case is one where skills are permanent, as in this case banks' profits would perfectly reveal banks' skills and fully dissipate the asymmetric information problem after the first period.) This, however, would be at the price of a lack of tractability of the model.

slowing down the accumulation of assets by the household and curbing the credit booms, e.g. through wealth effects in labor supply decisions or capital adjustment costs.[36] Overall, the typical SBC breaks out as productivity returns back to its average level after having remained above it for an unusually long time. This corresponds to a situation where the household has accumulated assets in anticipation of a fall in productivity that took more time than expected to materialize. In effect, one can show that the household over–accumulates assets, compared to what a central planner would, because she does not internalize the effects of her savings decisions on the equilibrium corporate loan rate and, ultimately, on the likelihood of a banking crisis.

Figure 10: Typical path (III)



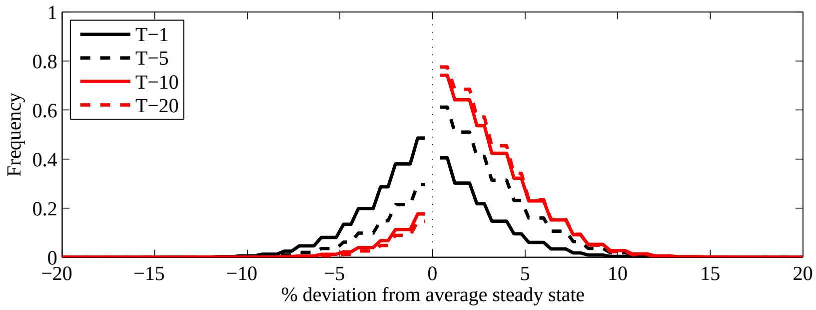
The above description of the typical path to crises, however, conceals the variety of banking crises, which the model features. In particular, SBCs may also be caused by an exogenous large negative TFP shock. For example, in Appendix C.1 we consider the response of the economy to such a large shock and show that, in this case, the crisis breaks out on impact. A 7% drop from the average steady state of TFP is sufficient to generate a crisis. To get a sense of the frequency of such shock–driven crises with respect to the typical credit boom–led ones, we report in Figure 11 the distributions of the deviations from their average levels of (the log of) TFP, the credit/output ratio, and a measure of financial imbalances in the kth

[36]On *(i)*: see Shin et al. (2011) for some evidence that substitution effects between corporate bonds and bank loans played an important role in mitigating the impact of the 2007–2008 crisis in the US, and De Fiore and Uhlig (2012) for a DSGE model with such effects. On *(ii)*: The difficulty of introducing such wealth effects in our model lies with the numerical resolution of the equilibrium, as in this case the threshold $\overline{a}_t$ would not depend exclusively on state variable z_t anymore.

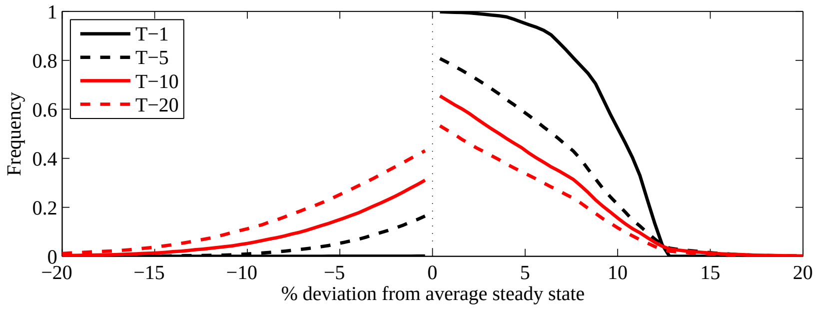
period before an SBC erupts, for $k = 1, 5, 10, 20$.

Figure 11: Distribution of percentage deviations before an SBC

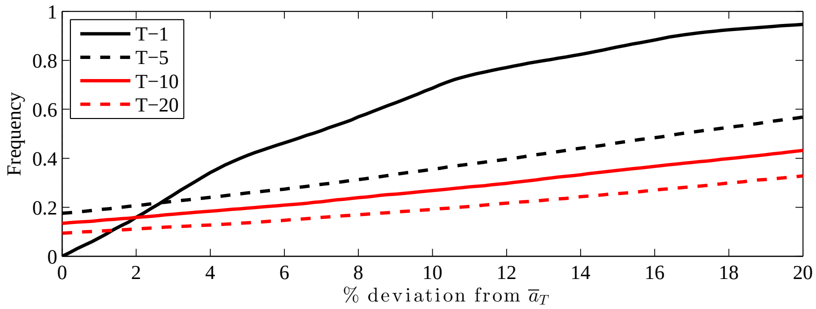
(a) TFP (in Logarithm)



(b) Credit/Output Ratio



(c) Financial Imbalances



Note: The figure above shows ergodic cumulative distributions. Panels (a) and (b) show $\mathbb{P}(\Delta x_{T-k} \geqslant \zeta)$ in the case of positive percentage deviations ζ (x-axis, right hand side) from average steady state, and $\mathbb{P}(\Delta x_{T-k} \leqslant -\zeta)$ in the case of negative deviations $-\zeta$ (x-axis, left hand side), where $\Delta x_{T-k} = \log(x_{T-k}) - \log(\overline{x})$, T is the date of the crisis, $\overline{x}$ is the average steady state of variable x calculated over our 500,000 simulations, and x is the logarithm of TFP or credit/output ratio. Panel (c) shows the probability that the level of assets k periods before the crisis be less than ζ% away from the absorption capacity of the banking sector, $\mathbb{P}(\Delta a_{T-k} \leqslant \zeta)$, with $\Delta a_{T-k} = \log(\overline{a}_{T-k}) - \log(a_{T-k})$.

Panel (a) shows that few SBCs occur after a year where productivity is significantly below its average steady state. (For instance only 10% of SBCs follow a year where TFP is 5% below average.) TFP is moreover almost symmetrically distributed around its average in the year before an SBC, which essentially means that the sign of TFP shocks —or more precisely: whether TFP is below or above its average— is not in itself informative about whether a SBC is about to break out. The skewed distributions of lagged TFP (for $k = 5, 10, 20$) at a medium to long term horizon even suggests that banking crises tend to be preceded by rather long periods of high —and not low— TFP. This is because prolonged sequences of TFP levels above average feed credit booms; and credit booms are conducive to SBCs. Panel (b) shows for instance that the credit to output ratio is 5% above its average steady state in the year that precedes the SBC in 95% of the SBCs. These results point to a general and important property of our model. Namely, the size of financial imbalances that prevail at the time shocks hit the economy are more relevant for the dynamics of the economy than the size and sign of the shocks themselves. Panel (c) illustrates this point more accurately; it shows the distributions of a measure of financial imbalances k periods before SBCs (again, for $k = 5, 10, 20$). We measure financial imbalances by the distance between banks' absorption capacity and banks' core liabilities, $\log(\overline{a}_t) - \log(a_t)$. The smaller this distance, the larger the imbalances, and the less resilient to adverse shocks the banking sector. Typically, a distance of less than, say, 5% reflects large financial imbalances. Panel (c) shows that 70% of the SBCs break out in the year after financial imbalances have reached this threshold. This result confirms that most crises in our model break out endogenously, without an adverse exogenous shock happening at the same time.

The above discussion prompts the question why, in our model, crises are more likely to occur in good, rather than in bad, times. This is due to the asymmetric effects of permanent income mechanisms on financial stability over the business cycle. Bad times in the model are typically times where TFP is low and the household dis–saves. By lowering the TFP threshold $\overline{z}_t$, below which a crisis breaks out, the fall in savings makes crises less likely. Hence, in bad times, the dynamics of savings tends to stabilize the financial sector. In good times, in contrast, TFP is high and the household accumulates assets, which by raising $\overline{z}_t$ makes crises more likely. In this case the dynamics of savings tends to destabilize the financial sector. This asymmetric effect of savings is the basic reason why credit–boom led crises are so prevalent in our model. Figure 16 in next section also illustrates this point.

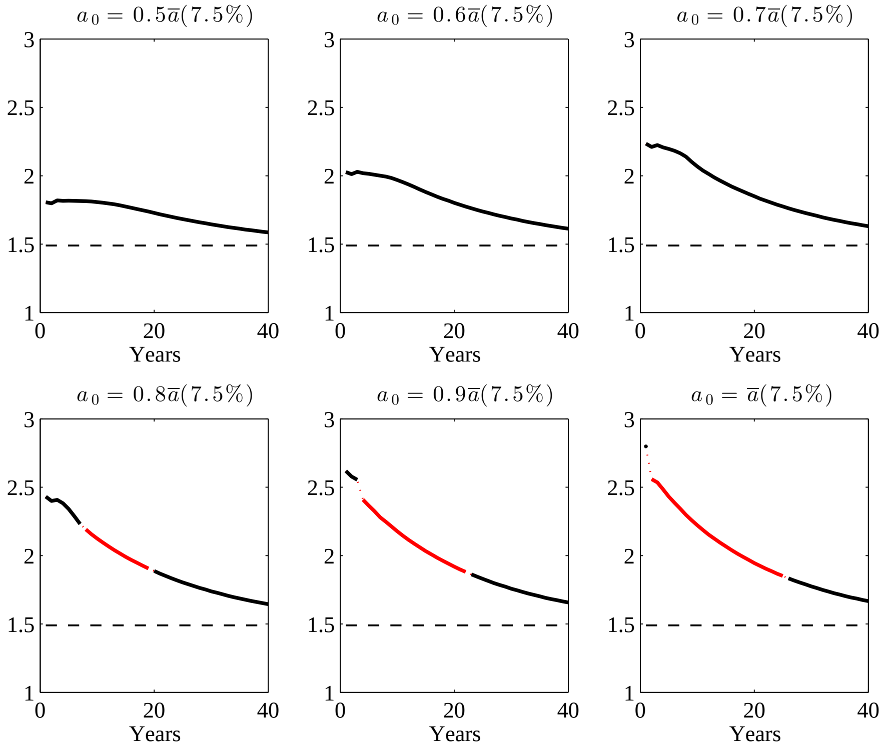
5.3 The Role of Financial Imbalances

This section focuses on the effects of financial imbalances on the frequency, the duration, and the amplitude of banking crises. As above, we measure financial imbalances as the distance between $\overline{a}_t$ and a_t. The experiment compares the transition dynamics of the economy toward its average steady state as the initial asset position of the agent varies. Building upon the previous experiment, which showed that crises are most likely to occur after a productivity boom, we impose that productivity is initially 7.5% above average in all the simulations. We simulate the economy starting from six different initial levels of assets ranging from 0% (large financial imbalances) to 50% (small imbalances) below absorption capacity.[37] Figures 12 and 13 report the median adjustment dynamics of output and the credit to assets ratio across

[37]We do not take a stance on the reasons why the economy reached these particular levels of assets in the first place, as this is not relevant here.

500,000 simulations.[38]

Figure 12: Output dynamics: Sensitivity to Initial Conditions



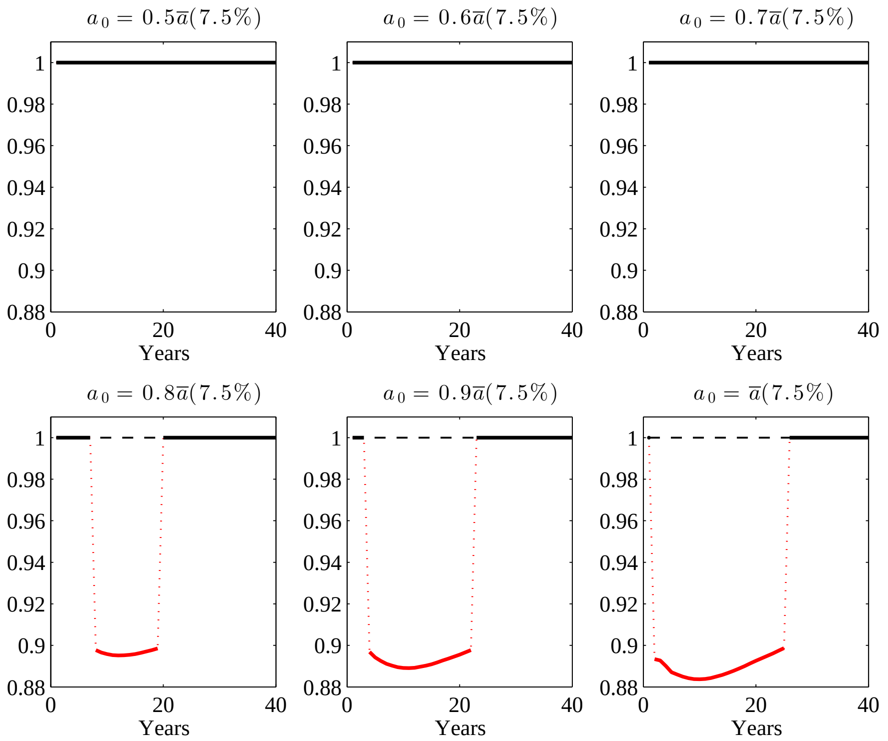
Dynamics in normal times, Dynamics in a systemic banking crisis, long–run average. $\overline{a}(7.5\%)$ denotes the banks' absorption when productivity is 7.5% above average

We find that the closer the banking sector to its absorption capacity, the more likely and the sooner the crisis. No crisis occurs so long the initial asset position of the economy is at least 30% below the absorption capacity. Financial imbalances pave the way to crises. Interestingly, initial conditions not only affect the probability and the timing of a crisis, but also its duration and amplitude. Figure 13 thus shows that the length of the credit crunch raises from 12 to 19 periods as the initial level of assets goes from 20% to 10% below absorption capacity. In that case the initial drop in output raises from 1.85% to 5.88% (see Figure 12).

Figure 14 illustrates the banking sector's resilience to shocks by reporting the proportion of paths featuring SBCs along the transition dynamics. In the economy with small imbalances (upper left panel), this proportion does not exceed 20% at any point in time. This suggests that, for this economy to experience a SBC, extreme TFP levels (and therefore extreme sequences of TFP shocks) in the 20% tail of the TFP distribution must occur. In contrast, the economy with the largest imbalances (lower right panel) is much more fragile and unlikely

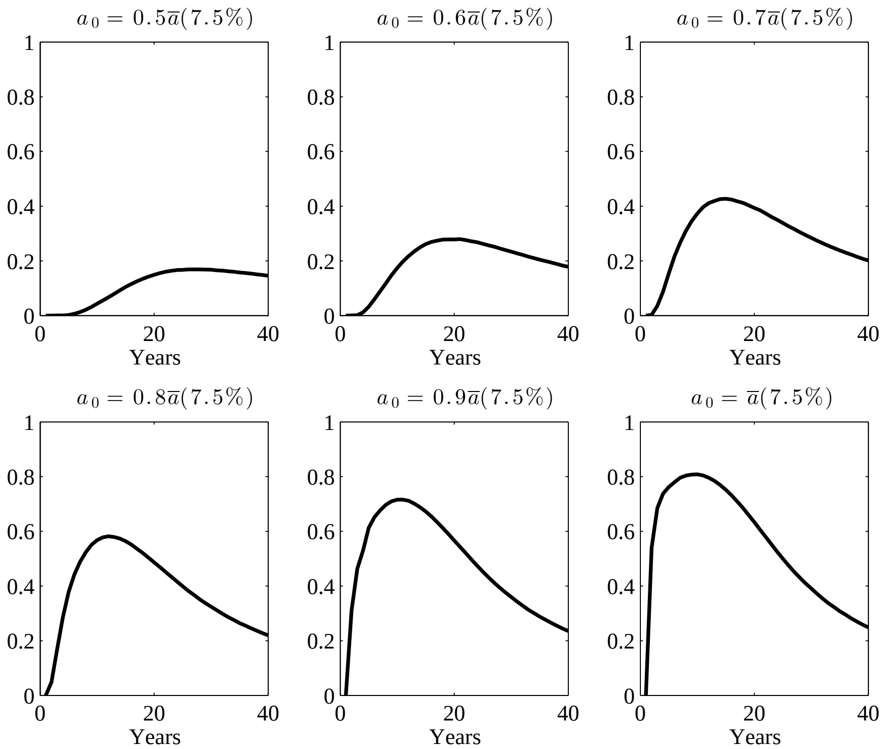
[38]The mean path would be the usual representation. However this representation would average out the effects of the financial crises, which do not break out along every one of the 500,000 simulated paths. This is the reason why, in the context of our model the median path is more informative.

Figure 13: Credit/Assets Dynamics: Sensitivity to Initial Conditions



Dynamics in normal times, Dynamics in a systemic banking crisis, long–run average. $\overline{a}(7.5\%)$ denotes the banks' absorption when productivity is 7.5% above average

Figure 14: Frequency of SBCs: Sensitivity to Initial Conditions



This figure reports the evolution of the frequency of SBCs during the transition toward the average steady state.

to escape a crisis, irrespective of the size of the shocks that hit it. In this economy more than 80% of the trajectories indeed feature a SBC.

The Role of Permanent Income Mechanisms: Financial imbalances are brought about by the household's rational expectations and permanent income behaviors —consumption smoothing and precautionary savings. As imbalances build up, the representative household realizes that the likelihood of a crisis increases, and accumulates assets to smooth her consumption profile. By doing so, however, the household inadvertently exacerbates the imbalances that may precipitate the crisis. To illustrate and assess quantitatively the importance of saving behaviors in the model, we replicated the above experiment using a "Solow" version of our model, where the saving rate is constant and the household neither smooths consumption nor accumulates precautionary savings.[39] We set the saving rate, s, such that the economy reaches the same average steady state as in our benchmark economy ($s = 0.22$). For comparison purposes, we also set the initial level of productivity 7.5% above average in all the simulations, and only let the initial level of assets vary. Because in this version of the model only large imbalances lead to a crisis, we only report in Figure 15 the results for initial levels of asset greater than $0.8\overline{a}(7.5\%)$.

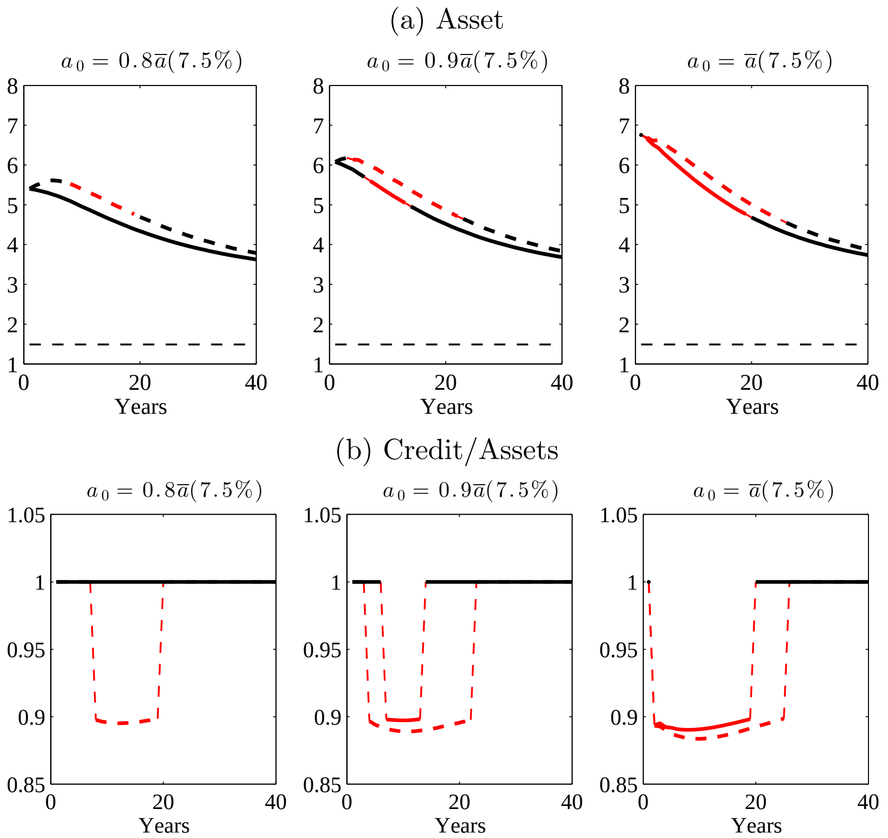
Panel (a) of Figure 15 shows that the household accumulates assets more slowly in the absence (plain line) than in the presence (dashed line) of permanent income mechanisms. The upshot is that SBCs are less frequent, shorter, and milder in the Solow version of the model. When the economy is started 10% below absorption capacity (middle panel), for instance, the crisis takes more time to materialize and entails a smaller credit crunch, with the credit to assets ratio falling by 10.3%, against 11.1% in the benchmark. In this case output drops by 3 percentage points less and the crisis is 11 years shorter than in the benchmark.

More generally, financial imbalances do not build up as much, and therefore do not play as big a role on financial stability in the absence of permanent income mechanisms. To illustrate this point, we derive the typical path to crisis in the Solow economy and find that the typical crisis is shock–driven, as opposed to credit–boom driven (compare Figures 16 and 8).[40] This result highlights the importance of saving dynamics to explain why, in our model, systemic banking crises break out in the midst of credit intensive booms (Fact #3).

[39] See the companion technical appendix for a detailed description of this version of the model.

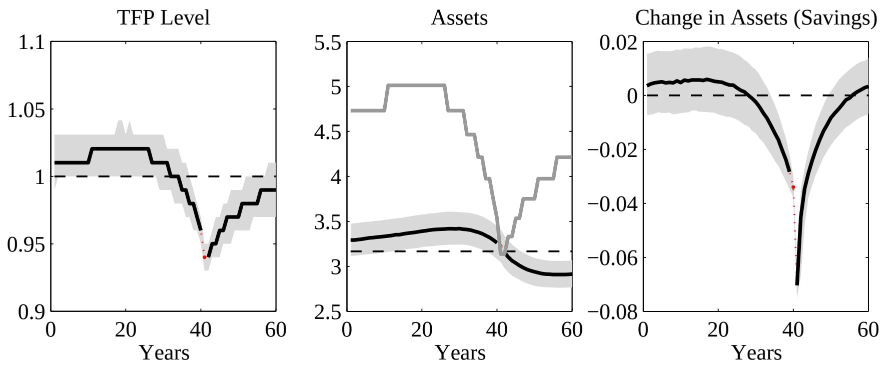
[40] This result is also reflected in the difference between the optimal decision rules in our benchmark model and those in the Solow model, which all cross the 45 degree line in the normal times regime. For a comparison between the optimal decision rules, see the companion technical appendix.

Figure 15: The Role of Permanent Income Mechanisms



(- - - -) Dynamics in normal times in the Solow version (Benchmark Model), (- - - -) Dynamics in a systemic banking crisis in the Solow version (Benchmark Model), - - - - long–run average. $\overline{a}(7.5\%)$ denotes the banks' absorption when productivity is 7.5% above average.

Figure 16: Typical path: Solow model



Dynamics in normal times, Dynamics in a systemic banking crisis, Dynamics of $\overline{a}_t$, - - - - long–run average, 66% Confidence Band.

5.4 Assessment of the Model

We now assess the ability of the model to account for the main stylized facts on systemic banking crises (see Section 2). As a first step, we simulate the model and generate long time series (500,000 periods) for output and credit. We then identify all recessions. A "recession" in the simulations starts when the economy experiences growth below a threshold that guarantees that, on average, the model generates recessions 10% of the time, as in the actual data (see Table 1). We further identify as financial recessions those recessions that feature a systemic banking crisis. Table 3 reports summary recession statistics similar to those reported in Table 1.

Table 3: Statistics on recessions and banking crises

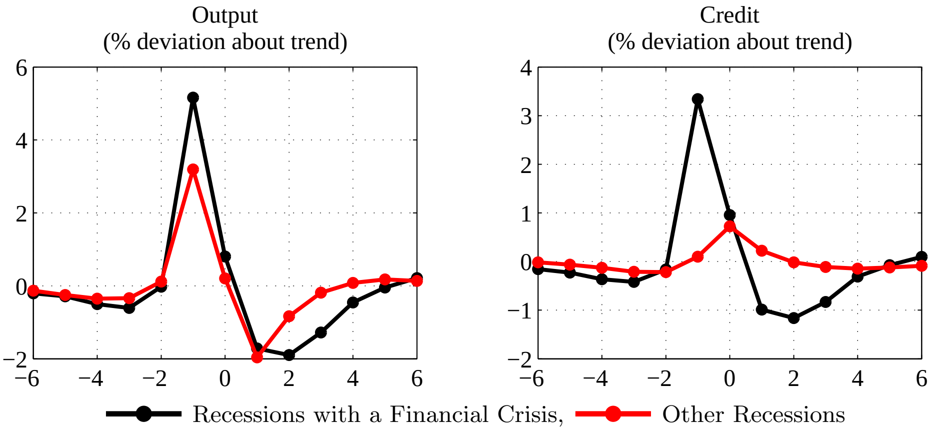
	Frequency (%)	Magnitude (%) from peak to trough	Duration (Years)
Systemic Banking Crises (SBC)	2.69	–	–
All recessions	10.00	12.08 (7.30)	2.08
Recessions with SBC	13.00	17.87 (10.50)	2.62
Recessions w/o SBC	87.00	10.04 (6.73)	1.90

Note: As in Table 1, the magnitudes reported into parentheses are calculated using the HP–filtered series of output, and are thereby corrected for the underlying trend in output. Notice that the average amplitude (duration) of financial recessions is slightly higher, with a 17.87% fall in output (2.62 years), than along the typical path in Figure 10, where it is around 15% (2 years). These differences are due to the non-linearities in the model, which imply that the average response of the economy to shocks (as reported above) differs from the response of the economy to the average shock (as reported in Figure 10).

On average, the duration of a recession featuring a banking crisis is 26% larger than that of other recessions. It also replicates the observation that recessions with SBCs are deeper than the average recession (48%). The model, however, overall over–estimates the amplitude and duration of recessions (compare Table 3 with Table 1). As discussed in section 5.2, this result is due to *(i)* our assumptions that firm financing is fully intermediated and *(ii)* the absence of wealth effects on labor supply decisions. Since the model was calibrated so as to obtain systemic banking crises every forty years in the simulations, the model replicates —by construction— Fact #1 (*i.e.* crises are rare events). Hence, we focus on Fact #2 and Fact #3.

Fact #2: Financial recessions are deeper and last longer than other recessions. Figure 17 replicates Figure C.6 and reports the average dynamics of the cyclical component of output and credit around financial recessions (dark line) and the other recessions (red line). The striking observation from the figure is its similarity with the actual data. The model is able to replicate the facts that financial recessions particularly deep and that both

Figure 17: Dynamics of GDP and Credit gaps around Recessions



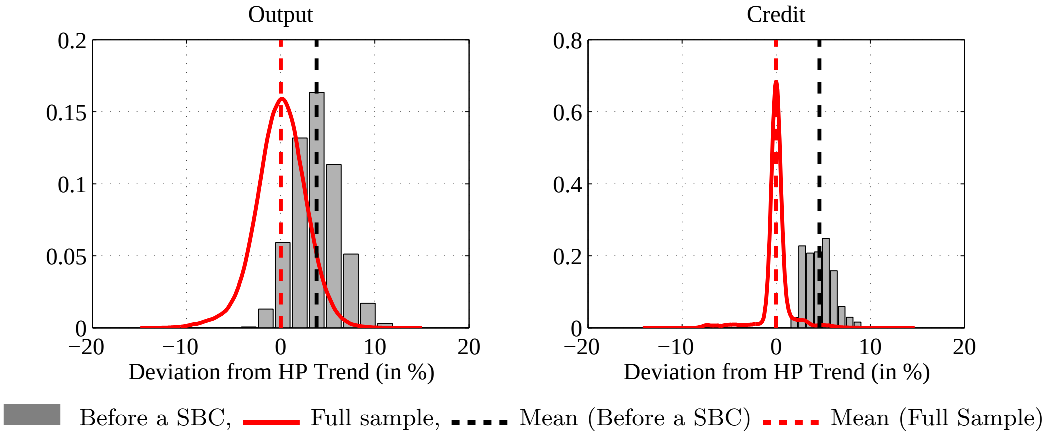
Note: For the sake of consistency, we treated the simulated series of output and credit in the same way as we treated the actual data in Section 2. Output and credit are first expressed in logarithm and the cyclical component is obtain using the HP–filter with $\lambda = 6.25$. Thus, as in Figure C.6, the reported % deviations are the average % deviations around a Hodrick–Prescott trend. The implied magnitude of financial recessions (the left chart) is about 7.1%, which is lower than that 10.50% reported in Table 3. This discrepancy reflects the fact that the statistics in Table 3 also take into account recessions of more than 6 years. We find similar results when we consider the % deviations of output and credit from their respective linear trends (see the companion technical appendix).

output and credit are much larger (5.2% and 3.4% above trend, respectively) before financial recessions than before regular recessions. In particular, as in the data, the credit gap drops to -1% during a SBC, much more than in a regular recession. The phase that follows the credit crunch also presents a U–shape pattern that reflects the greater persistence of financial recessions. Financial recessions last longer because they tend to occur in good times, when TFP is above mean and likely to go down. The exogenous fall in TFP that tends to follow upon the typical financial recession works to delay the recovery.

Fact #3: Systemic banking crises break out in the midst of credit intensive booms. Figure 18 replicates Figure 2 and reports the ergodic distributions of output and credit gaps in the full sample (red line) and in the year preceding a systemic banking crisis (histogram). This figure too bears a lot of similarities with its empirical counterpart. Both output and credit gaps are above trend in the year that precedes a banking crisis. One may argue that this result is not so surprising, since recessions (either in the data or in the model) should —by construction— follow expansions. However, we have seen with Figure 17 that the output and credit gaps are both much higher before a SBC than before a regular recession, and we find similar results for the distributions of GDP and credit growth rates (not reported). In other terms, these results are not driven by the way we filter the series.

As a last experiment, we want to illustrate the prediction properties of our model. To do so, we compute the probabilities that a SBC breaks out at a one– and a two–year horizon

Figure 18: Distributions of GDP and credit gaps

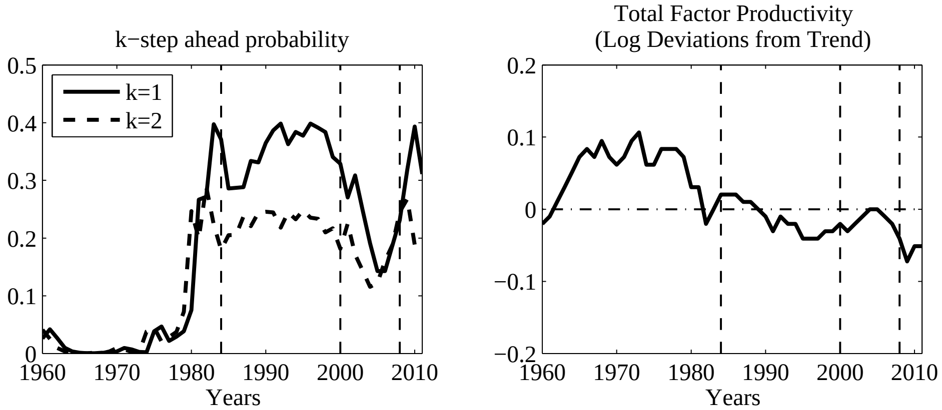


—as defined in Definition 2— and compare them with the actual US data over the period 1960-2011. The US experienced two banking crises during this period, both within the last three decades: the Savings & Loans crisis in 1984, and the subprime crisis in 2008. Jordà et al. (2011a) do not count the 2000 dotcom bubble bust as a *banking* crisis. We are interested in whether our model is able to predict these events. In the model, the crisis probability in year t (t=1960,. . . ,2011) is fully determined by the level of productivity, z_t, and the level of assets in the banking sector, a_t. The dynamics is initialized using the level of assets in the US banking sector in 1960, a_{1960}, taken from Jordà et al. (2011a). The model is then fed with Fernald (2009)'s series of total factor productivity, which we beforehand detrended in order to be consistent with our stationary setup,[41] and the model's internal dynamics are then used to compute the probabilities. Figure 19 reports the implied k–step ahead probability of a crisis for k=1,2 (left panel) as well as the underlying series of total factor productivity (right panel).

Considering its stylized character and the limited amount of information that it uses, the model predicts the US crisis episodes remarkably well. The one–step ahead probability (left panel) is essentially zero until the early 80s, spikes to 40% within three years, in 1981-1983, just before the Savings & Loans crisis, and remains relatively high throughout the rest of the period. While the probability is still 30% the year before the dotcom bubble bust, it goes down to 15% in 2001-2005 and picks up again as of 2007. At 20% the 2007 crisis probability is still relatively high by historical standards, and further increases in 2008-10, as the Great Recession materializes. The two–step ahead probability follows a similar pattern, with a one year lead. To a large extent, these results reflect the relationship in the model between financial stability and total factor productivity (right panel), which persistently stayed 10%

[41] More precisely, we linearly detrend the logarithm of Fernald's "naive" productivity series over the sample 1947–2011. We use the whole sample to use as much information as available. We then use the so–detrended series to construct the level of productivity, z_t, for the period 1960–2011. For the initial value of bank assets, a_{1960}, we apply the percentage deviation from linear trend of the banking sector's assets observed in 1960, to the model steady state associated with z_{1960}.

Figure 19: k–step ahead Probabilities of a Financial Crisis (k=1,2)



Note: The vertical thin dashed lines correspond to the 1984 Savings & Loans, the 2000 dotcom and 2008 crises.

above trend until the early 80s and has fluctuated below trend since then. We find similar results when we use Fernald's TFP series corrected for the rate of factor utilization, or when we detrend the series of TFP using a break in the trend to account for a structural US productivity slowdown in the mid 70s (see the companion technical appendix).

6 Discussion

6.1 Sensitivity Analysis

We now turn turn to the sensitivity of the properties of the model to the parameters. We simulate the model for 500,000 periods, compute the means of some key quantities across these simulations, and compare the results with our benchmark calibration (see Table 4).

Risk Averse Economies Are Prone to Crises: We first vary the utility curvature parameter σ from our benchmark 4.5 to values 2 and 10, therefore changing the degree of risk aversion of the household. By making the household more willing to accumulate assets for precautionary motives, *ceteris paribus*, the increase in σ works to raise the quantity of assets banks have to process without affecting banks' absorption capacity — leaving banks more exposed to adverse shocks. The probability of a crisis is thus higher than in the benchmark (5.4% *versus* 2.7%). In other words, the risk averse economy is paradoxically more prone to systemic banking crises. It also experiences deeper and longer crises than the benchmark economy, with output falling by 1.1 percentage point more from peak to trough and crises lasting 1.4 year longer. The main reason is that, by accumulating more assets, the economy builds up larger imbalances that make it difficult to escape crises once they occur. Accordingly, the banking sector of the risk averse economy is also less efficient, with an interest rate spread of 2.09%, against 1.71% in the benchmark. In contrast, less risk averse economies are

Table 4: Sensitivity Analysis

	Benchmark	σ		ν		θ	λ	σ_z	ρ_z
		2	10	0.25	1	0.20	35	0.02	0.95
Returns:									
Interbank (ρ)	0.86	1.04	0.23	0.84	0.83	0.40	1.34	0.89	0.72
Corporate (R)	4.35	4.55	3.70	4.28	4.41	5.50	3.70	4.32	4.29
deposit/equity (r)	2.64	2.96	1.61	2.52	2.80	2.61	2.67	2.55	2.59
Spread ($R - r$)	1.71	1.59	2.09	1.77	1.60	7 2.89	1.03	1.77	1.70
$\overline{R}$	2.43	2.43	2.43	2.43	2.43	4.83	0.41	2.43	2.43
Crises:									
Probability	2.69	1.20	5.43	3.31	0.99	7.34	0.16	3.35	1.90
Duration	2.62	1.89	4.08	2.90	1.97	5.06	1.87	2.82	2.92
Amplitude	17.87	17.30	19.00	19.94	11.96	16.90	15.80	19.36	16.08

Note: All numbers are averages over a long simulation of 500,000 periods and, except for durations, are expressed in percents. In the case where the persistence of the technology shock is raised to $\rho_z = 0.95$, the standard deviation of the innovation was rescaled so as to maintain constant the volatility of TFP.

more resilient to systemic banking crises. When $\sigma = 2$, for instance, SBCs occur only with a 1.20% probability.

Elastic Labor Supply Increases the Probability of a Crisis: We vary the (inverse of the) labor supply Frisch elasticity and consider values $\nu = 1/4$ and $\nu = 1$, the latter being rather standard in the DSGE literature (see Christiano et al., 2005). The Frisch elasticity affects the speed of accumulation of assets. In good times, for example, the household typically works more and, therefore, accumulates assets faster when $\nu = 1/4$ (high elasticity) than when $\nu = 1$ (low elasticity). Since financial imbalances build up faster in the high elasticity economy, this economy is also more prone to crises: when ν is reduced from 1 to 1/4, the probability of a crisis increases from 0.99% to 3.31%. Moreover, when the crisis breaks out the impact of the credit crunch on output is magnified by the larger response of hours worked; hence crises are also more severe and longer lasting in economies with higher labor elasticity.

Contract Enforceability and Bank Efficiency Improve Financial Stability: The third column of the table reports statistics for an economy where the cost of diversion is lower than in the benchmark. The increase in θ from 0.1 to 0.2 works, in the first place, to aggravate the moral hazard problem between banks. Banks have to deleverage in order to keep issuing interbank claims. Of course, in the general equilibrium there are some counter–balancing effects. For example, since the banking sector is less efficient the return on bank deposit/equity goes down, the household reduces her savings, banks reduce their supply of credit to the firm, and the equilibrium corporate loan rate tends to increase, which works to restore banks' incentives. But such effects are of second order. Overall, the crisis probability jumps from 2.69% to 7.34%.

A change in the distribution of banks has qualitatively similar effects. In the fourth column we consider an economy where $\lambda = 35$, implying a high concentration of the distribution of banks towards the top. The banking sector as a whole is more efficient than in the benchmark. Since efficient banks have lower incentives to divert cash, the moral hazard problem is less stringent and counterparty fears recede. Lenders tolerate a higher market funding ratio, aggregate demand and the interbank rate rise, which crowds the less efficient banks out of the demand side of the interbank loan market. As a result, the crisis probability drops from 2.7% to 0.16%. When they occur banking crises are however slightly more pronounced and more protracted. This is because they break out after a longer build–up phase, upon larger imbalances.

Higher Uncertainty is Conducive to Crises: As the only source of uncertainty, technology shocks play a crucial role in the model. In the last two columns in Table 4 we consider two experiments. First, we increase the volatility of the shock by raising σ_z from 0.018 to 0.02, leaving its persistence unchanged. The consequences are straightforward: the household accumulates more assets for precautionary motives, the corporate loan rate decreases with respect to benchmark, and the financial sector is more fragile. Next, we increase the persistence of the shock by raising ρ_z shock from 0.9 to 0.95, leaving its volatility unchanged. This change exerts two opposite effects on financial stability. On the one hand, following a positive productivity shock the household does not accumulate assets as fast as in the benchmark. Hence financial imbalances build up more slowly. On the other hand, the productivity level is more likely to remain high for longer, implying that imbalances accumulate over a longer period of time and, therefore, grow larger. In our calibration, the first effect dominates, and the probability of a crisis overall decreases.

6.2 Bank Leverage, Bank Defaults

One caveat of our model is that it abstracts from the dynamics of bank leverage and bank defaults, two elements yet important from a macro–prudential perspective. The reason is that, absent any friction between the banks and the household, the deposit to equity ratio is indeterminate. In this section we introduce a friction that permits us to pin this ratio down and to analyze the evolution of bank leverage and bank defaults along the typical path to crisis. For simplicity, we will keep a framework where the bank deposit to equity ratio does not have any impact on aggregate dynamics, so that the typical path to crisis is unchanged. To derive this ratio we make two additional assumptions. First, we assume that banks can collect deposits, which we define as the risk–free asset of the economy. Since deposits must be risk–free, there is a limit to the quantity of deposits that a given bank can collect. This limit is related to the value of the bank's assets in the worst possible state of the nature, *i.e.* γa_{t+1} for a bank born in t. (The worst possible state of the nature for a bank is one where the bank is inefficient and there is a crisis.) Second, we assume that banks are run by risk–neutral managers, whose objective is to maximize their banks' expected return on equity. That managers do not discount risk —or have a different discount factor than the shareholder's— is a standard assumption that creates a wedge between the managers and the shareholder's objectives. While in equilibrium the shareholder will be indifferent between equity and deposits, managers will in contrast be willing to raise as much deposits as possible. To model defaults, we further assume that the shareholder has a *partial* liability, in the sense that she must recapitalize the banks that are unable to repay their deposits up to a fraction ξ of the banks' assets, with $\xi > 0$. The shareholder's liability thus amounts to ξa_{t+1}. Under this assumption, a manager is able to raise deposits up to $(\gamma + \xi)a_{t+1}$, which is more than what her bank will be able to repay, should a bad enough state of the nature materialize. We will say in this case that the bank defaults.[42] Let r^d_{t+1} be the risk–free (non–state contingent) gross return on deposits d_{t+1} and $r^e_{t+1}(p)$ be the *ex post* return on bank p's equity e_{t+1} at the end of period $t+1$ with, by definition, $a_{t+1} \equiv d_{t+1} + e_{t+1}$ and

$$r^e_{t+1}(p) \equiv \frac{r_{t+1}(p)a_{t+1} - r^d_{t+1}d_{t+1}}{e_{t+1}}, \tag{10}$$

where $r_{t+1}(p)$ is defined in (5). Since bank p defaults when $r^e_{t+1}(p) < 0$ the type $\widetilde{p}_{t+1}$ of the marginal bank that defaults at the end of period $t+1$ is given by

$$\widetilde{p}_{t+1} = r^{-1}_{t+1}\left(r^d_{t+1}\frac{d_{t+1}}{a_{t+1}}\right)\mathbb{I}_{\overline{p}_{t+1}\leqslant r^{-1}_{t+1}\left(r^d_{t+1}\frac{d_{t+1}}{a_{t+1}}\right)}, \tag{11}$$

where $\mathbb{I}_{\overline{p}_{t+1}\leqslant r^{-1}_{t+1}\left(r^d_{t+1}\frac{d_{t+1}}{a_{t+1}}\right)}$ is an indicator function that takes value one if $\overline{p}_{t+1} \leqslant r^{-1}_{t+1}\left(r^d_{t+1}\frac{d_{t+1}}{a_{t+1}}\right)$ and zero otherwise. The household's no–arbitrage condition between deposits and equity then requires that

$$\mathbb{E}_t(u'(c_{t+1})r^e_{t+1}) = r^d_{t+1}\mathbb{E}_t(u'(c_{t+1})), \tag{12}$$

where $r^e_{t+1} \equiv \int_0^1 r^e_{t+1}(p)\mathrm{d}\mu(p)$. This condition relies on r^d_{t+1} being a risk–free rate determined at the end of period t, yields r^d_{t+1} so that the household is indifferent between deposits and

[42] An alternative way to introduce bank defaults could be to assume that a deposit insurance scheme is in place, that makes it possible for the managers to raise more risk–free deposits. The difference with our setup is that deposit insurance is a subsidy to the banking sector, which would then have to be accounted for in the Euler equation (4). In the companion technical appendix of the paper, we show that deposit insurance would affect the dynamics of savings and the typical path to crisis, but only marginally. The constraint that a given bank's deposits are limited by the bank's equity capital (see constraint 24 below) has indeed very little impact on the aggregate dynamics in our model. The reason is that banks have the possibility to issue equity on the market so as to raise their deposits and, ultimately, increase their supply of corporate loans. That is, bank capital is not scarce. In this respect, our model stands in contrast with other models, like Gertler and Kiyotaki (2009), where banks are not allowed to issue equity. This restriction is very severe because it implies banks can only accumulate net worth through retained earnings. And in these models banks typically never live long enough to accumulate enough net worth out of financing constraints. As a result, bank capital is scarce and, by limiting the quantity of credit that banks can supply, it affects the aggregate dynamics. This is not the case here.

equity in equilibrium, and implies that $r^d_{t+1} < \mathbb{E}_t(r^e_{t+1})$. Risk–neutral managers choose the deposit to equity ratio so as to maximize their respective banks' expected returns on equity[43]

$$\max_{d_{t+1}/e_{t+1}} \beta\mathbb{E}_t(r_{t+1}) + \beta\left(\mathbb{E}_t(r_{t+1}) - r^d_{t+1}\right)\frac{d_{t+1}}{e_{t+1}} \tag{13}$$

subject to the feasibility constraint that deposits must be risk–free,

$$d_{t+1}r^d_{t+1} \leqslant (\gamma+\xi)a_{t+1}. \tag{14}$$

Since (from (19) and (21)) $\mathbb{E}_t\left(r_{t+1}\right) > r^d_{t+1}$ it is clear that managers want to raise as much deposits as possible and that, in equilibrium, constraint (24) binds. Hence,

$$\frac{d_{t+1}}{e_{t+1}} = \frac{\gamma+\xi}{\max\left(r^d_{t+1}-\gamma-\xi, 0^+\right)}. \tag{15}$$

One can now easily derive borrowing (lending) banks' optimal leverage ℓ^b_{t+1} (ℓ^l_{t+1}) as

$$\ell^b_{t+1} = \frac{d_{t+1}+\phi_{t+1}(d_{t+1}+e_{t+1})}{e_{t+1}} \text{ and } \ell^l_{t+1} = \frac{d_{t+1}}{e_{t+1}}, \tag{16}$$

as well as the mass Δ_{t+1} of the banks that default as

$$\Delta_{t+1} = \mu\left(\widetilde{p}_{t+1}\right). \tag{17}$$

Bank leverage depends on both the market funding ratio and the deposit to equity ratio. It is easy to see from Proposition 1 and relation (26) that these two ratios move in opposite directions. On the one hand the market funding ratio increases with ρ_{t+1}. On the other hand, the deposit to equity ratio decreases with r^d_{t+1}, as bank managers can afford raising more deposits when deposits become cheaper. Since ρ_{t+1} and r^d_{t+1} move together, it is not clear how leverage moves along the typical path. To investigate this, we set $\xi = 0.4\%$ so that in the simulations 1% of the banks default on average every year; with this calibration bank leverage is 16.13 on average (*i.e.* the bank capital ratio is around 6%), no bank defaults in normal times, and 12.5% of the banks default in crisis times (see Table 5). The evolution of

Table 5: Leverage and Default

Leverage			Default		
Borrowing Banks	Lending Banks	Aggregate	Normal	Crisis	Average
19.82	11.41	16.13	0	12.55%	0.90%

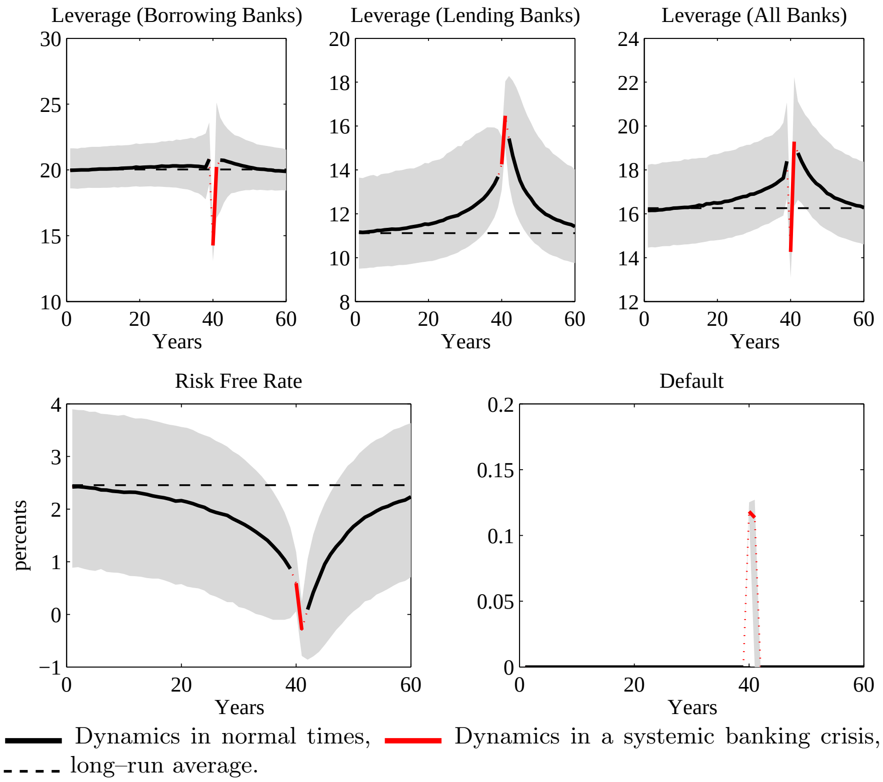
Note: All numbers are averages over a long simulation of 500,000 periods.

leverage, defaults, and the deposit rate along the typical path to crisis is reported in Figure 20.[44] In the run–up to the typical crisis, bank leverage overall increases for all banks, be

[43]More precisely, bank managers choose d_{t+1}/e_{t+1} to maximize $\beta\mathbb{E}_t\left(r^e_{t+1}\right)$ subject to the identities $r_{t+1}a_{t+1} \equiv r^d_{t+1}d_{t+1} + r^e_{t+1}e_{t+1}$ and $a_{t+1} \equiv d_{t+1}+e_{t+1}$. If managers' incentives were aligned with the banks' shareholder, then they would instead maximize $\beta\mathbb{E}_t\left(\frac{u'(c_{t+1})}{u'(c_t)}r^e_{t+1}\right)$ and be indifferent between equity and deposits.

[44]We focus on these variables because, the household being indifferent between equity and deposits, the deposit to equity ratio does not have any effect on the macroeconomic dynamics and all the other variables of the model follow the same path as in Figures 8–10.

Figure 20: Typical Path: Leverage and Default



they lenders or borrowers. For borrowers, in particular, the rise in the deposit to equity ratio offsets the diminution of the market funding ratio (see Figure 9); hence, even for those banks leverage is slightly pro–cyclical in normal times. During the crisis, in contrast, the evolution of leverage differs across banks, depending on whether or not they rely on market funding. While the leverage of borrowing banks (which finance themselves through the market) falls during the crisis, that of lending banks (which only finance themselves through equity and deposits) increases. The reason is that, in our model, bank equity is more volatile than retail deposits but less volatile than interbank loans. This result highlights the importance of the (un)stability of banks' various sources of funding in determining the dynamics of leverage. Finally, the lower right panel in Figure 20 shows that a significant mass of banks need to be recapitalized during systemic banking crises.

7 Conclusion

We develop a simple macroeconomic model that accounts for key features of systemic banking crises and helps understand why, historically, banking crises *(i)* occur in unusually good times and *(ii)* bring about the most severe recessions. The primary cause of systemic banking crises in the model is the accumulation of assets by the household in anticipation of future adverse shocks. When long enough, this asset accumulation process generates financial imbalances (credit boom, disproportionately large bank balance sheets), which, in the best case, lower the resilience of the banking sector to shocks and, in the worst case, lead endogenously to a crisis. In this context, adverse random shocks themselves play a secondary role, as what matters the most is the conditions under which these shocks occur. This last result is of particular importance from a policy making perspective: it implies that systemic banking crises are predictable. We show indeed that, although very stylized, our model provides with a simple tool to detect financial imbalances and predict future crises.

References

Bernanke, B., M. Gertler and S. Gilchrist, 1999, "The Financial Accelerator in a Quantitative Business Cycle Framework," Handbook of Macroeconomics, Volume 1, Chapter 11, J..B. Taylor and M. Woodford eds, Elsevier.

Bianchi, J., 2011, "Overborrowing and Systemic Externalities in the Business Cycle," American Economic Review, 101(7), 3400–3426.

Bianchi, J. and E. G. Mendoza, 2010, "Overborrowing, Financial Crises and 'Macroprudential' Policy," NBER Working Paper No. 16091.

Borio, C. and M. Drehmann, 2009, "Assesing the risk of banking crises: Revisited", BIS Quarterly Review, March, 29–46.

Borio, C. and P. Lowe, 2002, "Assessing the risk of banking crises", BIS Quarterly Review, December, 43–54.

Brunnermeier, M. and L. Pedersen, 2009, "Market Liquidity and Funding Liquidity," Review of Financial Studies, 22 (6): 2201–2238.

Brunnermeier, M. and Y. Sannikov, 2012, "A Macroeconomic Model with a Financial Sector," mimeo, Princeton University.

Burkart, M. and T. Ellingsen, 2004, "In-Kind Finance: A Theory of Trade Credit,", American Economic Review, 94(3): 569–590.

Christiano, L., M. Eichenbaum and C. Evans, 2005,, "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," Journal of Political Economy, 113(1), 1–45.

Claessens, S., A. Kose and M. Terrones, 2008, "How Do Business and Financial Cycles Interact?," IMF working Paper No. WP/11/88.

Claessens, S., A. Kose and M. Terrones, 2011, "What Happens During Recessions, Crunches and Busts?," IMF working Paper No. WP/08/274.

Cooper, R. and A. John, 1988, "Coordinating Coordination Failures in Keynesian Models," Quarterly Journal of Economics, (103)3, August, 441–463.

Cooper, R., D. DeJong, R. Forsythe, and T. Ross, 1990, "Selection Criteria in Coordination Games: Some Experimental Results," American Economic Review, vol. 80, No. 1, March, 218–233.

De Fiore, F. and H. Uhlig, 2012, "Corporate Debt Structure and the Financial Crisis," mimeo, European Central Bank.

Fernald, J., 2009, "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," mimeo, Federal Reserve Bank of San Francisco.

Gertler, M. and P. Karadi, 2011, "A Model of Unconventional Monetary Policy', Journal of Monetary Economics, forthcoming.

Gertler, M. and N. Kiyotaki, 2009, "Financial Intermediation and Credit Policy in Business Cycle Analysis," Mimeo, Princeton University.

Gertler, M. and N. Kiyotaki, 2012, "Banking, Liquidity and Bank Runs in an Infinite Horizon Economy," Mimeo, Princeton University.

Gorton, G., 1988,“Banking Panics and Business Cycles,” Oxford Economic Papers, New Series, vol. 40, No. 4, 751–781.

Gorton, G., 2010, Interview with “The Region,” December, downloadable from `http://www.minneapolisfed.org/pubs/region/10-12/gorton.pdf`.

Gorton, G., 2012, “Some Reflections on the Recent Financial Crisis,” NBER Working Paper No 18397.

Greenwood, J., Z. Hercowitz and G. Huffman, 1988,“Investment, Capacity Utilization, and the Real Business Cycle,” American Economic Review, vol 78, No. 3, 402–417.

Hart, O., 1995, “Firms, contracts, and Financial Structure,” Oxford University Press.

Hahm, J–H., H–S. Shin and K. Shin, 2011, “Non-Core Bank Liabilities and Financial Vulnerability,” mimeo, Princeton University.

He, Z. and A. Krishnamurthy, 2012, “A Macroeconomic Framework for Quantifying Systemic Risk,” mimeo, University of Chicago.

Holmström, B. and J. Tirole, 1997, “Financial Intermediation, Loanable Funds, and the Real Sector,” Quarterly Journal of Economics, vol. CXII, No 3, August, 663–691.

Jermann, U. and V. Quadrini, 2010, “Macroeconomic Effects of Financial Shocks”, American Economic Review, forthcoming.

Jordà, O., M. Schularick and A. Taylor, 2011a, “Financial Crises, Credit Booms and External Imbalances: 140 Years of Lessons,” IMF Economic Review, Vol. 59, 340–378.

Jordà, O., M. Schularick and A. Taylor, 2011b, “When Credit Bites Back: Leverage, Business Cycles, and Crises,” NBER Working Paper No.17621.

Judd, K. L. 1992, “Projection methods for solving aggregate growth models,” Journal of Economic Theory, Vol. 58, 410-52.

Korinek, A., 2010, “Systemic Risk-Taking: Accelerator Effects, Externalities, and Regulatory Responses,” mimeo, University of Maryland.

Laeven, L. and F. Valencia, 2008, “Systemic Banking Crises: A New Database,” IMF Working Paper No 08224.

Maddaloni, A. and J–L. Peydro, 2011, “Bank Risk-Taking, Securitization, Supervision, and Low Interest Rates: Evidence from the Euro Area and US Lending Standards,” Review of Financial Studies, Vol. 24, 2121–65, 2011.

Mehra, R. and E. Prescott, 1985, “The Equity Premium: A Puzzle,” Journal of Monetary Economics, 145–161.

Mendoza, E., 2010, “Sudden Stops, Financial Crises and Leverage,” American Economic Review, Vol. 100(5), 1941–66.

Mendoza, E. and K., Smith, 2005, “Quantitative Implications of a Debt-Deflation Theory of Sudden Stops and Asset Prices,” Journal of International Economics, Vol. 70(1), 82–114.

Minsky, H., 1977, “A Theory of Systemic Fragility,” in E. Altman and A. Sametz, eds., Financial Crises.

Ravn, M. and H. Uhlig, 2002, “On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations,” Review of Economics and Statistics, May, Vol. 84, No. 2, 371–376.

Reinhart, C. and K. Rogoff, 2009, "This Time is Different — Eight Centuries of Financial Folly," Princeton University Press.

Schularick, M. and A. Taylor, 2011, "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008," American Economic Review, Vol. 102, 1029–1061.

Shin, H–S., 2008, "Securitisation and Financial Stability," mimeo, Princeton University.

Shin, H–S., 2010, "Risk and Liquidity," Clarendon Lecture in Finance, Oxford University Press.

Shin, H–S., and K. Shin., 2011, "pro–cyclicality and Monetary Aggregates," NBER Working Paper No 16836.

Stiglitz, J. and A. Weiss, 1981, "Credit Rationing in Markets with Imperfect Information," American Economic Review, vol. 71, No. 3 (June), 393–410.

Tauchen, G. and R. Hussey, 1991, "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," Econometrica, vol. 59, No. 2, 371–96.

Taylor, J., 2009, "The Financial Crisis and the Policy Responses:An Empirical Analysis of What Went Wrong," Critical Review, vol. 21, 341–364.

Tirole, J., 2006, "The Theory of Corporate Finance," Princeton University Press.

Uhlig, H., 2010, "A Model of Systemic Bank Run," Journal of Monetary Economics, vol. 57, 78–96.

— APPENDIX —

A Proofs of Propositions

Proposition 1: The program of a borrowing bank writes

$$\max_{\phi_t} \; pR_t(1+\phi_t) - \rho_t\phi_t$$
$$s.t.\; pR_t(1+\phi_t) - \rho_t\phi_t \geqslant \rho_t$$
$$\gamma(1+\theta\phi_t) \leqslant \rho_t$$

The participation constraint indicates that only banks with ability $p \geqslant \overline{p}_t \equiv \rho_t/R_t$ will borrow. We focus on this segment of the market. The problem simplifies to

$$\max_{\phi_t} pR_t(1+\phi_t) - \rho_t\phi_t$$
$$s.t.\; \gamma(1+\theta\phi_t) \leqslant \rho_t$$

for $p \geqslant \overline{p}_t$. Let us denote by λ the Lagrange multiplier associated to the incentive constraint, the first order conditions are then

$$pR_t - \rho_t = \gamma\theta\lambda$$
$$\lambda(\rho_t - \gamma(1+\theta\phi_t)) = 0$$

It follows that $\lambda > 0$ for all $p > \overline{p}_t$. The result follows. ■

Proposition 2: The market clearing condition (9a) in normal times together with the optimal demand for capital yields the normal times equilibrium corporate loan rate $R_t = f_k(a_t; z_t) + 1 - \delta$. The interbank market is at work if and only if this rate is above $\overline{R}$, *i.e.*, if and only if $a_t \leqslant \overline{a_t} \equiv f_k^{-1}(\overline{R} + \delta - 1; z_t)$. ■

Proposition 3: The proof proceeds in two steps.

i) From (9) and Proposition 2, we show that the corporate loan rate increases during a credit crunch:

$$\lim_{a_t \nearrow \overline{a}_t} R_t + \delta - 1 = \lim_{a_t \nearrow \overline{a}_t} f_k(a_t; z_t) = \lim_{a_t \searrow \overline{a}_t} f_k(a_t; z_t)$$
$$< \lim_{a_t \searrow \overline{a}_t} f_k\left((1 - \mu(\gamma/R_t))\, a_t; z_t\right) = \lim_{a_t \searrow \overline{a}_t} R_t + \delta - 1$$

ii) From (7), we know that $\lim_{a_t \nearrow \overline{a}_t} r_t/R_t = \int_{\overline{p}_t}^1 p \frac{\mathrm{d}\mu(p)}{1-\mu(\overline{p}_t)}$, which conditional expectation monotonically increases in $\overline{p}_t$. Using point (i) and the inequality $\rho_t \geqslant \gamma$, we have that $\lim_{a_t \nearrow \overline{a}_t} \overline{p}_t > \lim_{a_t \nearrow \overline{a}_t} \gamma/R_t > \lim_{a_t \searrow \overline{a}_t} \gamma/R_t$. Hence, $\lim_{a_t \nearrow \overline{a}_t} r_t/R_t > \lim_{a_t \searrow \overline{a}_t} \int_{\gamma/R_t}^1 p \frac{\mathrm{d}\mu(p)}{1-\mu(\gamma/R_t)}$. Finally, we know from (7) that

$$\lim_{a_t \searrow \overline{a}_t} r_t/R_t = \lim_{a_t \searrow \overline{a}_t} \frac{\gamma}{R_t}\mu\left(\frac{\gamma}{R_t}\right) + \int_{\gamma/R_t}^1 p\,\mathrm{d}\mu(p) < \lim_{a_t \searrow \overline{a}_t} \int_{\gamma/R_t}^1 p \frac{\mathrm{d}\mu(p)}{1-\mu(\gamma/R_t)},$$

which implies $\lim_{a_t \nearrow \overline{a}_t} r_t/R_t > \lim_{a_t \searrow \overline{a}_t} r_t/R_t$. The result follows. ■

B Equations of the Model with endogenous Labour Supply

1. $y_t = z_t k_t^{\alpha} h_t^{1-\alpha} + (\gamma + \delta - 1)(a_t - k_t)$
2. $R_t = \alpha k_t^{\frac{-\upsilon(1-\alpha)}{\upsilon+\alpha}} z_t^{\frac{1+\upsilon}{\upsilon+\alpha}} \left(\frac{1-\alpha}{\vartheta}\right)^{\frac{1-\alpha}{\upsilon+\alpha}} + 1 - \delta$
3. $\left(c_t - \vartheta \frac{h_t^{1+\upsilon}}{1+\upsilon}\right)^{-\sigma} = \beta \mathbb{E}_t \left[\left(c_{t+1} - \vartheta \frac{h_{t+1}^{1+\upsilon}}{1+\upsilon}\right)^{-\sigma} r_{t+1} \right]$
4. $h_t = \left(\frac{(1-\alpha) z_t}{\vartheta}\right)^{\frac{1}{\upsilon+\alpha}} k_t^{\frac{\alpha}{\upsilon+\alpha}}$
5. $\overline{a}_t \equiv ((1-\alpha)/\vartheta)^{\frac{1}{\upsilon}} \left(\alpha / \left(\overline{R} + \delta - 1\right)\right)^{\frac{\upsilon+\alpha}{\upsilon(1-\alpha)}} z_t^{\frac{1+\upsilon}{\upsilon(1-\alpha)}}$
6. $i_t = a_{t+1} - (1-\delta) a_t$

If $a_t \leqslant \overline{a}_t$ (normal times)

7*a*. $k_t = a_t$

8*a*. $\frac{r_t}{R_t} = \int_{\overline{p}_t}^{1} p \frac{\mathrm{d}\mu(p)}{1-\mu(\overline{p}_t)}$

9*a*. $\overline{p}_t = \frac{\rho_t}{R_t}$

10*a*. $R_t = \frac{\rho_t}{\mu^{-1}\left(\frac{\rho_t - \gamma}{\rho_t - (1-\theta)\gamma}\right)}$, with $\rho_t > \overline{\rho}$

11*a*. $y_t = c_t + i_t + (R_t - r_t) a_t$

If $a_t > \overline{a}_t$ (crisis times)

7*b*. $k_t = a_t - \mu(\gamma/R_t) a_t$

8*b*. $\frac{r_t}{R_t} = \frac{\gamma}{R_t}\mu(\gamma/R_t) + \int_{\gamma/R_t}^{1} p \, \mathrm{d}\mu(p)$

9*b*. $\overline{p}_t = \gamma/R_t$

10*b*. $\rho_t = \gamma$

11*b*. $y_t = c_t + i_t + (R_t - r_t) a_t - (R_t - \gamma)(a_t - k_t)$

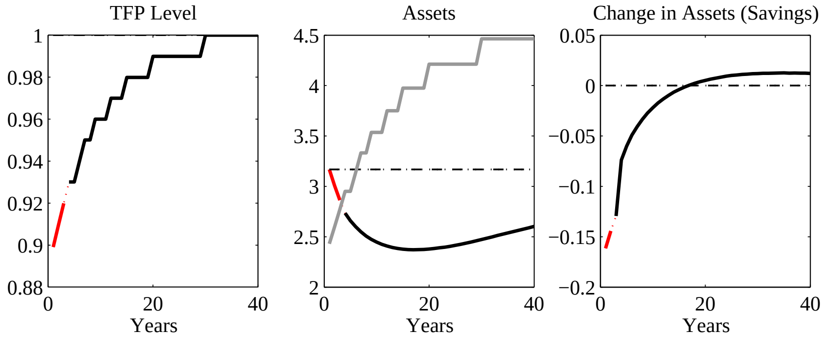
Few comments are in order here. *(i)* In autarky assets $a_t - k_t$ are stored for return γ. Since capital depreciates at rate δ, the value added of storage is $\gamma + \delta - 1$, as reflected in equation 1. (Remember that $\gamma + \delta - 1 > 0$.) *(ii)* The good market clearing conditions 11a and 11b are derived from Walras' law and the agents' budget constraints. Summing up the household and the firm's budget constraints (1) and (3), one gets: $c_t + a_{t+1} = z_t k_t^{\alpha} h_t^{1-\alpha} + (1-\delta) k_t + r_t a_t - R_t k_t$, which simplifies to $c_t + i_t = y_t - (\gamma+\delta-1)(a_t - k_t) - (1-\delta)(k_t - a_t) + r_t a_t - R_t k_t \Longleftrightarrow y_t = c_t + i_t + (R_t - r_t) a_t - (R_t - \gamma)(a_t - k_t)$. *(iii)* The last two terms of this accounting identity correspond to the aggregate cost of financial frictions. In autarky, for instance, this cost is, by definition, equal to $\int_{\gamma/R_t}^{1} (1-p) R_t a_t \, \mathrm{d}\mu(p)$, which indeed simplifies to (using 7b and 8b): $(1 - \mu(\gamma/R_t)) R_t a_t - a_t \int_{\gamma/R_t}^{1} p R_t \, \mathrm{d}\mu(p) = R_t k_t - r_t a_t + \gamma(k_t - a_t) = (R_t - r_t) a_t - (R_t - \gamma)(a_t - k_t)$.

C Additional Figures

C.1 Response to a Large Negative Shock

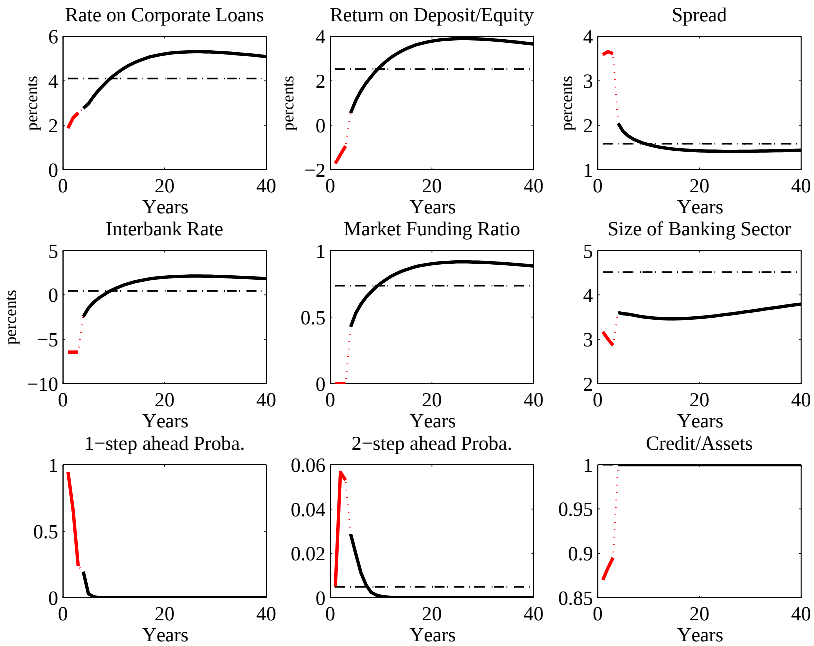
The economy is started from the average steady state and is subjected to a 10% drop in total factor productivity in the initial period. As in other DSGE models featuring financial frictions, the model is able to generate a financial crisis from a large negative shock, provided that the shock is large enough. In our case, a 7% drop from average steady state is sufficient to generate crises.

Figure C.1: Response to a large drop in TFP (I)



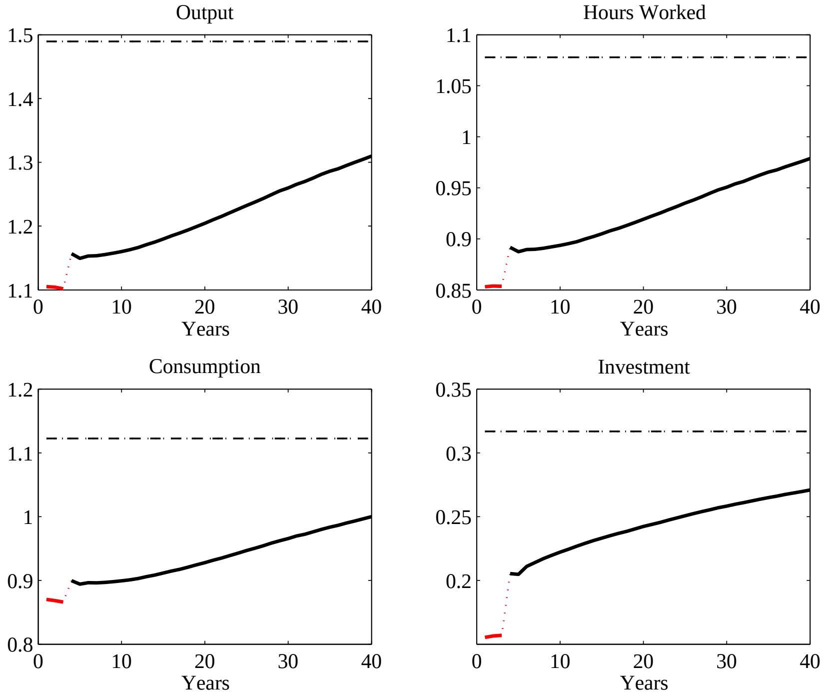
Dynamics in normal times, Dynamics in a systemic banking crisis, Dynamics of $\bar{a}_t$, long–run average.

Figure C.2: Response to a large drop in TFP (II)



Dynamics in normal times, Dynamics in a systemic banking crisis, long–run average.

Figure C.3: Response to a large drop in TFP (III)

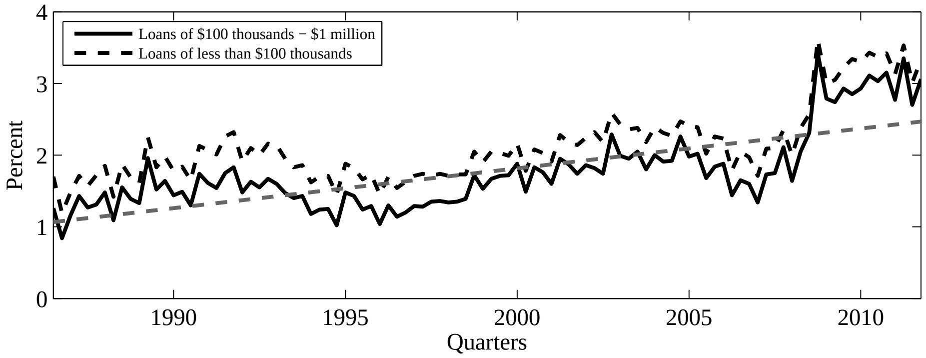


Dynamics in normal times, Dynamics in a systemic banking crisis, long–run average.

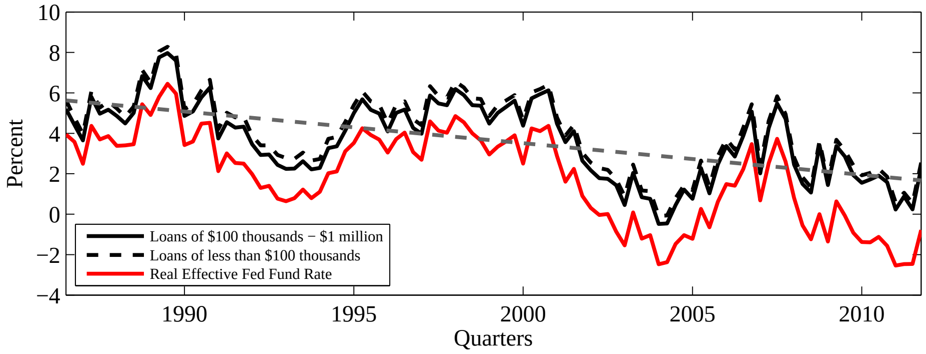
C.2 Evolution of US Interest Rates and Spreads

Figure C.4: Evolution of Various Corporate Loan Spreads

(a) Spread: Corporate loan rates - Federal Fund Rate



(b) Underlying Real Corporate Loan Rates



Trend line.

The effective federal fund rate is from the Federal Reserve Economic Database (`http://research.stlouisfed.org/fred2/series/FEDFUNDS?cid=118`). The corporate loan rates are from the Federal reserve (`http://www.federalreserve.gov/releases/e2/e2chart.htm`). Nominal rates are deflated using the annualized quarter on quarter rate of growth of the GDP implicit deflator (`http://research.stlouisfed.org/fred2/series/GDPDEF?cid=21`).

TECHNICAL APPENDIX

NOT INTENDED FOR PUBLICATION

A Comparative statics

This section reports some useful comparative statics that shed light on the dynamics at work in the model.

Result 1 *The corporate loan rate is a decreasing function of the capital stock, and an increasing function of the technology shock.*

This can be easily seen from the optimal behavior of the firm which states

$$R_t = f_k(k_t; z_t) + 1 - \delta$$

The result follows from the existence of marginal decreasing returns to scale with respect to the capital stock and the fact that the technology shock affects positively the production efficiency.

Proposition 4 *In a stable trade equilibrium*[45]*, the interbank market increases with the corporate loan rate*

$$\frac{\mathrm{d}\rho_t}{\mathrm{d}R_t} > 0$$

An increase in the corporate loan rate makes it more attractive for banks to borrow on the interbank market and lend to the firms. This implies that more banks are willing to borrow (positive effect on the extensive margin) and that each individual bank is willing to leverage more (positive effect on the intensive margin, See corollary 1).

Proof: The interbank market can be easily obtained from the market clearing condition

$$G(\rho_t, R_t) = \left(1 - \mu\left(\frac{\rho_t}{R_t}\right)\right)\phi_t - \mu\left(\frac{\rho_t}{R_t}\right) = 0 \tag{18}$$

where $G(\rho_t, R_t)$ is the net demand function for loans on the interbank market. Note that we restrict our attention to the stable Walrasian equilibrium. A sufficient condition for the equilibrium to be stable in the sense of Walrasian tatônnement is

$$\frac{\partial G_(\rho_t; R)}{\partial \rho_t} < 0 \Longleftrightarrow \mu'\left(\frac{\rho_t}{R_t}\right)\frac{1+\phi_t}{R_t} - \left(1 - \mu\left(\frac{\rho_t}{R_t}\right)\right)\frac{\mathrm{d}\phi_t}{\mathrm{d}\rho_t} > 0$$

Total differentiation of the market clearing condition then yields

$$\underbrace{\left[\mu'\left(\frac{\rho_t}{R_t}\right)\frac{1+\phi_t}{R_t} - \left(1 - \mu\left(\frac{\rho_t}{R_t}\right)\right)\frac{\mathrm{d}\phi_t}{\mathrm{d}\rho_t}\right]}_{>0}\mathrm{d}\rho_t = \underbrace{\mu'\left(\frac{\rho_t}{R_t}\right)\frac{1+\phi_t}{R_t}\frac{\rho_t}{R_t}}_{>0}\mathrm{d}R_t$$

Q.E.D. ■

[45] By this we mean an equilibrium in which trade takes place on the interbank market, such that there exists an equilirbium which is stable in the sense of a Walrasian tatônement process.

Corollary 1 *In a stable trade equilibrium, the individual interbank funding ratio increases with the corporate loan rate.*

Proof: Recall that, from the incentive constraint, we have

$$\phi_t = \frac{\rho_t - \gamma}{\theta\gamma} \iff \frac{\mathrm{d}\phi_t}{\mathrm{d}R_t} = \frac{1}{\gamma\theta}\frac{\mathrm{d}\rho_t}{\mathrm{d}R_t} > 0$$

Q.E.D. ■

Proposition 5 *In a stable trade equilibrium, the cutoff $\overline{p}_t$ increases with the corporate loan rate*

$$\frac{\mathrm{d}\overline{p}_t}{\mathrm{d}R_t} > 0$$

Proof: The cutoff is obtained from the participation constraint as $\overline{p}_t = \rho_t/R_t$, which implies that

$$\frac{\mathrm{d}\overline{p}_t}{\mathrm{d}R_t} = \frac{1}{R_t}\frac{\mathrm{d}\rho_t}{\mathrm{d}R_t} - \frac{\rho_t}{R_t^2}$$

Using proposition 4, this rewrites

$$\frac{\mathrm{d}\overline{p}_t}{\mathrm{d}R_t} = \frac{\rho_t}{R_t^2}\left(\frac{\mu'\left(\frac{\rho_t}{R_t}\right)\frac{1+\phi_t}{R_t}}{\mu'\left(\frac{\rho_t}{R_t}\right)\frac{1+\phi_t}{R_t} - \left(1-\mu\left(\frac{\rho_t}{R_t}\right)\right)\frac{\mathrm{d}\phi_t}{\mathrm{d}\rho_t}} - 1\right) = \frac{\rho_t}{R_t^2}\left(\frac{\left(1-\mu\left(\frac{\rho_t}{R_t}\right)\right)\frac{\mathrm{d}\phi_t}{\mathrm{d}\rho_t}}{\mu'\left(\frac{\rho_t}{R_t}\right)\frac{1+\phi_t}{R_t} - \left(1-\mu\left(\frac{\rho_t}{R_t}\right)\right)\frac{\mathrm{d}\phi_t}{\mathrm{d}\rho_t}}\right)$$

Using Corollary 1, we have $\mathrm{d}\phi_t/\mathrm{d}R_t > 0$ and hence $\mathrm{d}\overline{p}_t/\mathrm{d}R_t > 0$.
Q.E.D. ■

Proposition 6 *In a stable trade equilibrium, the rate on equity/deposit increases with the corporate loan rate*

$$\frac{\mathrm{d}r_t}{\mathrm{d}R_t} > 0$$

Proof: In normal times, the rate on deposits is given by

$$r_t = \frac{R_t}{1-\mu(\overline{p}_t)}\int_{\overline{p}_t}^1 p\mathrm{d}\mu(p)$$

Hence

$$\frac{\mathrm{d}r_t}{\mathrm{d}R_t} = \frac{\int_{\overline{p}_t}^1 p\mathrm{d}\mu(p)}{1-\mu(\overline{p}_t)} + \frac{\mu'(\overline{p}_t)\int_{\overline{p}_t}^1 p\mathrm{d}\mu(p) - \overline{p}_t\mu'(\overline{p}_t)(1-\mu(\overline{p}_t))}{(1-\mu(\overline{p}_t))^2}\frac{\mathrm{d}\overline{p}_t}{\mathrm{d}R_t}$$

Using the definition of r_t, this rewrites

$$\frac{\mathrm{d}r_t}{\mathrm{d}R_t} = \frac{r_t}{R_t} + \frac{\mu'(\overline{p}_t)}{1-\mu(\overline{p}_t)}\left(\frac{r_t}{R_t} - \overline{p}_t\right)$$

Since $\overline{p}_t = \rho_t/R_t$, we have

$$\frac{\mathrm{d}r_t}{\mathrm{d}R_t} = \frac{r_t}{R_t} + \frac{\mu'(\overline{p}_t)}{1-\mu(\overline{p}_t)}\left(\frac{r_t-\rho_t}{R_t}\right)\frac{\mathrm{d}\overline{p}_t}{\mathrm{d}R_t}$$

Given that[46] $r_t > \rho_t$ and $\mathrm{d}\overline{p}_t/\mathrm{d}R_t > 0$, the result follows.
Q.E.D. ■

[46]Recall that $r_t(p) = pR_t(1+\phi_t) - \rho_t\phi_t$. Let us define $p = \overline{p}_t + \varepsilon$, with $\varepsilon > 0$, then $r_t(p) = \rho_t + \varepsilon R_t(1+\phi_t) > \rho_t$. Since r_t is obtained from the mass of bank with ability $p > \overline{p}_t$, we have $r_t > \rho_t$.

Lemma 1 *In a stable trade equilibrium, the spread between the corporate loan and equity/deposit rates increases with the cutoff.*

Proof: In normal times, a measure of the spread is given by

$$\Psi_t \equiv \frac{R_t}{r_t} = \frac{1-\mu(\overline{p}_t)}{\int_{\overline{p}_t}^1 p\mathrm{d}\mu(p)}$$

Hence, totally differentiating and using the definition of Ψ_t

$$\frac{\mathrm{d}\Psi_t}{\mathrm{d}\overline{p}_t} = \frac{\mu'(\overline{p}_t)}{\int_{\overline{p}_t}^1 p\mathrm{d}\mu(p)}\left(\frac{R_t}{r_t}\overline{p}_t - 1\right) = \frac{\mu'(\overline{p}_t)}{\int_{\overline{p}_t}^1 p\mathrm{d}\mu(p)}\left(\frac{\rho_t}{r_t} - 1\right) < 0$$

Q.E.D. ■

Corollary 2 *In a stable trade equilibrium, we have*

$$\frac{\mathrm{d}\rho_t}{\mathrm{d}a_t} < 0, \quad \frac{\mathrm{d}r_t}{\mathrm{d}a_t} < 0, \quad \frac{\mathrm{d}\phi_t}{\mathrm{d}a_t} < 0, \quad \frac{\mathrm{d}\overline{p}_t}{\mathrm{d}a_t} < 0, \quad \frac{\mathrm{d}\Psi_t}{\mathrm{d}a_t} > 0,$$

$$\frac{\mathrm{d}\rho_t}{\mathrm{d}z_t} > 0, \quad \frac{\mathrm{d}r_t}{\mathrm{d}z_t} > 0, \quad \frac{\mathrm{d}\phi_t}{\mathrm{d}z_t} > 0, \quad \frac{\mathrm{d}\overline{p}_t}{\mathrm{d}z_t} > 0, \quad \frac{\mathrm{d}\Psi_t}{\mathrm{d}z_t} < 0$$

Proof: This result follows from the fact that in normal times, $a_t = k_t$ and from Result 1 in the following propositions, corollaries and lemmas.
Q.E.D. ■

This corollary tells us that, in normal times, the various interest rates in the economy decrease as agents accumulate assets, as $k_t = a_t$. Positive technology shocks balance the effects of accumulation.

Proposition 7 *During an SBC credit increases with the level of assets*

$$\frac{\mathrm{d}k_t}{\mathrm{d}a_t} > 0$$

Proof: In an SBC, we have $k_t = a_t\left(1-\mu\left(\frac{\gamma}{R_t}\right)\right)$. Total differentiation of the last expression yields

$$\mathrm{d}k_t = \left(1-\mu\left(\frac{\gamma}{R_t}\right)\right)\mathrm{d}a_t + \frac{\gamma}{R_t^2}a_t\mu'\left(\frac{\gamma}{R_t}\right)\underbrace{\frac{\partial R_t}{\partial k_t}}_{<0}\mathrm{d}k_t$$

Hence

$$\frac{\mathrm{d}k_t}{\mathrm{d}a_t} = \frac{1-\mu\left(\frac{\gamma}{R_t}\right)}{1-\frac{\gamma}{R_t^2}a_t\mu'\left(\frac{\gamma}{R_t}\right)\frac{\partial R_t}{\partial k_t}} \geqslant 0$$

Q.E.D. ■

Proposition 8 *During an SBC credit, the rate on equity increases with the corporate loan rate*

$$\frac{\mathrm{d}r_t}{\mathrm{d}R_t} > 0$$

Proof: In an SBC, the rate on equity is given by

$$r_t = \gamma\mu\left(\frac{\gamma}{R_t}\right) + R_t\int_{\frac{\gamma}{R_t}}^{1} p\,\mathrm{d}\mu(p)$$

Totally differentiating this last expression, we get

$$\begin{aligned}\mathrm{d}r_t &= -\left(\frac{\gamma}{R_t}\right)^2 \mu'\left(\frac{\gamma}{R_t}\right)\mathrm{d}R_t + \int_{\frac{\gamma}{R_t}}^{1} p\,\mathrm{d}\mu(p)\mathrm{d}R_t + \left(\frac{\gamma}{R_t}\right)^2 \mu'\left(\frac{\gamma}{R_t}\right)\mathrm{d}R_t \\ &= \int_{\frac{\gamma}{R_t}}^{1} p\,\mathrm{d}\mu(p)\mathrm{d}R_t\end{aligned}$$

Hence

$$\frac{\mathrm{d}r_t}{\mathrm{d}R_t} = \int_{\frac{\gamma}{R_t}}^{1} p\,\mathrm{d}\mu(p) \geqslant 0$$

Q.E.D. ■

Proposition 9 *During an SBC credit, the spread, Ψ_t, between rate on corporate loans and the rate on equity increases with the corporate loan rate*

$$\frac{\mathrm{d}\Psi_t}{\mathrm{d}R_t} > 0$$

Proof: We have

$$\Psi_t = \frac{R_t}{r_t} \Longrightarrow \mathrm{d}\Psi_t = \frac{\mathrm{d}R_t}{r_t} - \frac{R_t}{r_t^2}\mathrm{d}r_t$$

such that

$$\mathrm{d}\Psi_t = \frac{\mathrm{d}R_t}{r_t}\left(1 - \frac{R_t}{r_t}\frac{\mathrm{d}r_t}{\mathrm{d}R_t}\right) = \frac{\mathrm{d}R_t}{r_t}\left(1 - \frac{R_t}{r_t}\int_{\frac{\gamma}{R_t}}^{1} p\,\mathrm{d}\mu(p)\right)$$

From the expression of the equity rate, we have

$$\int_{\frac{\gamma}{R_t}}^{1} p\,\mathrm{d}\mu(p) = \frac{r_t}{R_t} - \frac{\gamma}{R_t}\mu\left(\frac{\gamma}{R_t}\right)$$

such that

$$\frac{\mathrm{d}\Psi_t}{\mathrm{d}R_t} = \frac{\gamma}{r_t}\mu\left(\frac{\gamma}{R_t}\right) > 0$$

Q.E.D. ■

Corollary 3 *During an SBC, we have*

$$\frac{\mathrm{d}\rho_t}{\mathrm{d}a_t} = 0, \quad \frac{\mathrm{d}r_t}{\mathrm{d}a_t} < 0, \quad \frac{\mathrm{d}\phi_t}{\mathrm{d}a_t} = 0, \quad \frac{\mathrm{d}\Psi_t}{\mathrm{d}a_t} < 0,$$

$$\frac{\mathrm{d}\rho_t}{\mathrm{d}z_t} = 0, \quad \frac{\mathrm{d}r_t}{\mathrm{d}z_t} > 0, \quad \frac{\mathrm{d}\phi_t}{\mathrm{d}z_t} = 0, \quad \frac{\mathrm{d}\Psi_t}{\mathrm{d}z_t} > 0$$

Proof: Results on ρ_t and ϕ_t originate in the collapse of the interbank market. Results on r_t and ψ_t are direct consequences of Propositions 7, 8, 9 and Result 1.
Q.E.D. ■

B Solution Method

The model is solved using a collocation technique. We first discretize the distribution of the technology shock using the approach proposed by Tauchen and Hussey, 1991 (See Section B.1). This leads to a Markov chain representation of the technology shock with $z_t \in \{z_1, \ldots, z_{n_z}\}$ and transition matrix $\Pi = (\pi_{ij})_{i,j=1}^{n_z}$ where $\pi_{ij} = \mathbb{P}(z_{t+1} = z_j | z_t = z_i)$. We use $n_z = 31$ in the sequel. This leads us to look for n_z decision rules $a_{t+1}(a_t; z_i)$. Since we have two regimes in the model —normal times and crises— we approximate the decision rules as

$$a_{t+1}(a_t; z_i) = \exp\left(\sum_{j=0}^{q} \xi_j^{\mathrm{N}}(z_i) T_j(\varphi(a_t))\right) \mathbb{I}_{a_t \leqslant \overline{a}_t} + \exp\left(\sum_{j=0}^{q} \xi_j^{\mathrm{C}}(z_i) T_j(\varphi(a_t))\right) \mathbb{I}_{a_t > \overline{a}_t},$$

where z_i denotes a particular level of the total factor productivity in the grid. $T_j(\cdot)$ is a Chebychev polynomial of order j, $\xi_j^{\mathrm{N}}(z_i)$ (resp. $\xi_j^{\mathrm{C}}(z_i)$) is the coefficient associated to this polynomial when the economy is in normal times (resp. in a systemic banking crisis) for the value of productivity z_i. $\varphi(a_t)$ is a function that maps the level of assets into the interval (-1,1). Finally $\mathbb{I}_{a_t \leqslant \overline{a}_t}$ (resp. $\mathbb{I}_{a_t > \overline{a}_t}$) is an indicator function which takes value one in normal (resp. crisis) times and 0 otherwise. The optimal decision rule $a_{t+1}(a_t; z_t)$ is given by the fixed point solution to the Euler equation.

More precisely, the algorithm proceeds as follows.

1. Choose a domain $[a_m, a_s]$ of approximation for a_t. We consider values that guarantee that the conditional steady state associated with each level of the technology shock can be reached. This led us to use $a_m = 0.5$ and $a_s = 20$.

2. Choose an order of approximation q, compute the $q+1$ roots of the Chebychev polynomial of order $q+1$ as

$$\zeta_k = \cos\left(\frac{(2k-1)\pi}{2(q+1)}\right) \text{ for } k = 1, \ldots, q+1$$

   and formulate an initial guess for $\xi^\ell(z_i)$ for $\ell = \{N, C\}$ and $i = 1, \ldots, n_z$. We chose $q = 15$ so as to obtain an accurate approximation of the decision rule.

3. Compute a_k, $k = 1, \ldots, q+1$ as

$$a_k = \begin{cases} \exp\left(\log(a_m) + (\zeta_k + 1)\frac{\log(\overline{a}(z_i)) - \log(a_m)}{2}\right) & \text{for normal times} \\ \exp\left(\log(\overline{a}(z_i)) + (\zeta_k + 1)\frac{\log(a_s) - \log(\overline{a}(z_i))}{2}\right) & \text{for crisis times} \end{cases}$$

   for $k = 1, \ldots, q+1$. Note that the grid of values for a depends fundamentally on the level of the technology shock as the threshold $\overline{a}(z_i)$ depends on the level of productivity where

$$\overline{a}(z_i) = \Gamma z_i^{\frac{1+\upsilon}{\upsilon(1-\alpha)}}, \text{ with } \Gamma \equiv \left(\frac{1-\alpha}{\vartheta}\right)^{\frac{1}{\upsilon}} \left(\frac{\alpha}{\overline{R} + \delta - 1}\right)^{\frac{\upsilon+\alpha}{\upsilon(1-\alpha)}}.$$

4. Using a candidate solution $\xi(z_i) = (\xi^{\mathrm{N}}(z_i), \xi^{\mathrm{C}}(z_i))$, compute

$$a_{t+1}(a_k; z_i) = \exp\left(\sum_{j=0}^{q} \xi_j^{\mathrm{N}}(z_i) T_j(\zeta_k)\right) \mathbb{I}_{a_k \leqslant \overline{a}_t} + \exp\left(\sum_{j=0}^{q} \xi_j^{\mathrm{C}}(z_i) T_j(\zeta_k)\right) \mathbb{I}_{a_k > \overline{a}_t}$$

   for each level of a_k, $k = 1, \ldots, q+1$ and the over quantities of the model using the definition of the general equilibrium of the economy. In particular obtain hours worked, $h_t(a_k, z_i)$ and household's income $e_t(a_k, z_i)$.

5. Use the values $a_{t+1}(a_k; z_i)$ and the approximation to obtain $a_{t+2}(a_k, z_i, z_j') = a_{t+2}(a_{t+1}(a_k; z_i), z_j')$. Solve the general equilibrium to obtain next period hours worked, $h_{t+1}(a_k, z_i, z_j')$, income, $e_{t+1}(a_k, z_i, z_j')$, and interest rate, $r_{t+1}((a_k, z_i, z_j')$.

6. Given that $c_t = e_t - a_{t+1}$, the Euler equation can be written as

$$\begin{aligned}a_{t+1}(a_k,z_i) =& e_t(a_k,z_i) - \vartheta\frac{h_t(a_k,z_i)^{1+\nu}}{1+\nu}\\ &-\left(\beta\sum_{j=1}^{n_z}\pi_{ij}\left(e_{t+1}(a_k,z_i,z_j') - a_{t+1}(a_k,z_i,z_j') - \vartheta\frac{h_{t+1}(a_k,z_i,z_j')^{1+\nu}}{1+\nu}\right)^{-\sigma}\right)^{-\frac{1}{\sigma}}\end{aligned}$$

at each node $k = 1,\dots,q+1$ and for all values of the technology shock, $i = 1,\dots,n_z$.

7. If all residuals are close enough to zero in the sense that their norm is less than $\varepsilon \equiv$`1e-6`, then stop, else use the right hand side of the Euler equation to obtain a new function $a_{t+1}(a_k,z_i)$, obtain new values for $\widehat{\xi}(z_i) = (\widehat{\xi}^{\mathrm{N}}(z_i), \widehat{\xi}^{\mathrm{C}}(z_i))$ using a collocation regression and update the candidate solution as

$$\xi'(z_i) = \lambda\xi(z_i) + (1-\lambda)\widehat{\xi}(z_i)$$

B.1 Tauchen–Hussey (1991)

Tauchen and Hussey (1991) provide a simple way to discretize an AR(1) processf the form

$$z_{t+1} = \rho z_t + \varepsilon_{t+1}$$

where $\varepsilon_{t+1} \rightsquigarrow \mathscr{N}(0,\sigma^2)$. This implies that

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2}\left(\frac{z_{t+1}-\rho z_t}{\sigma}\right)^2\right\}\mathrm{d}z_{t+1} = \int f(z_{t+1}|z_t)\mathrm{d}z_{t+1} = 1$$

which illustrates the fact that s is a continuous random variable. Tauchen and Hussey propose to replace the integral by

$$\int\Phi(z_{t+1};z_t,\overline{s})f(z_{t+1}|\overline{s})\mathrm{d}z_{t+1} \equiv \int\frac{f(z_{t+1}|z_t)}{f(z_{t+1}|\overline{s})}f(z_{t+1}|\overline{s})\mathrm{d}z_{t+1} = 1$$

where $f(z_{t+1}|\overline{s})$ denotes the density of z_{t+1} conditional on the fact that $z_t = \overline{s}$ (in fact the unconditional density function), which in our case implies that

$$\Phi(z_{t+1};z_t,\overline{s}) \equiv \frac{f(z_{t+1}|z_t)}{f(z_{t+1}|\overline{s})} = \exp\left\{-\frac{1}{2}\left[\left(\frac{z_{t+1}-\rho z_t}{\sigma}\right)^2 - \left(\frac{z_{t+1}}{\sigma}\right)^2\right]\right\}$$

then we can use the standard linear transformation and impose $\zeta_t = z_t/(\sigma\sqrt{2})$ to get

$$\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp\left\{-\left((\zeta_{t+1}-\rho\zeta_t)^2-\zeta_{t+1}^2\right)\right\}\exp\left(-\zeta_{t+1}^2\right)\mathrm{d}\zeta_{t+1}$$

for which we can use a Gauss–Hermite quadrature. Given the quadrature nodes ζ_i and weights ω_i, $i = 1,\dots,n$, the quadrature leads to the formula

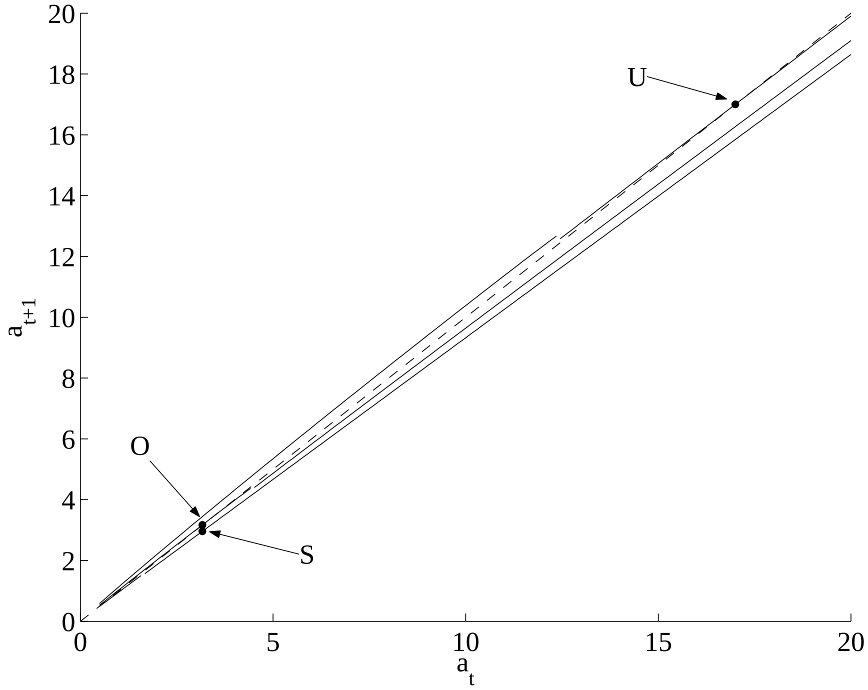
$$\frac{1}{\sqrt{\pi}}\sum_{j=1}^{n}\omega_j\Phi(\zeta_j;\zeta_i;\overline{x}) \simeq 1$$

and the quantity $\omega_j\Phi(\zeta_j;\zeta_i;\overline{x})/\sqrt{\pi}$ is an "estimate" $\pi_{ij} \equiv \mathrm{Prob}(z_{t+1} = z_j|z_t = z_i)$ of the transition probability from state i to state j. Since, in general, $\sum_{j=1}^{n}\pi_{ij} = 1$ will not hold exactly. Tauchen and Hussey therefore propose the following modification:

$$\pi_{ij} = \frac{\omega_j\Phi(z_j;z_i;\overline{s})}{\sqrt{\pi}\varsigma_i}$$

where $\varsigma_i = \frac{1}{\sqrt{\pi}}\sum_{j=1}^{n}\omega_j\Phi(z_j;z_i;\overline{x})$.

Figure B.5: Decision Rules

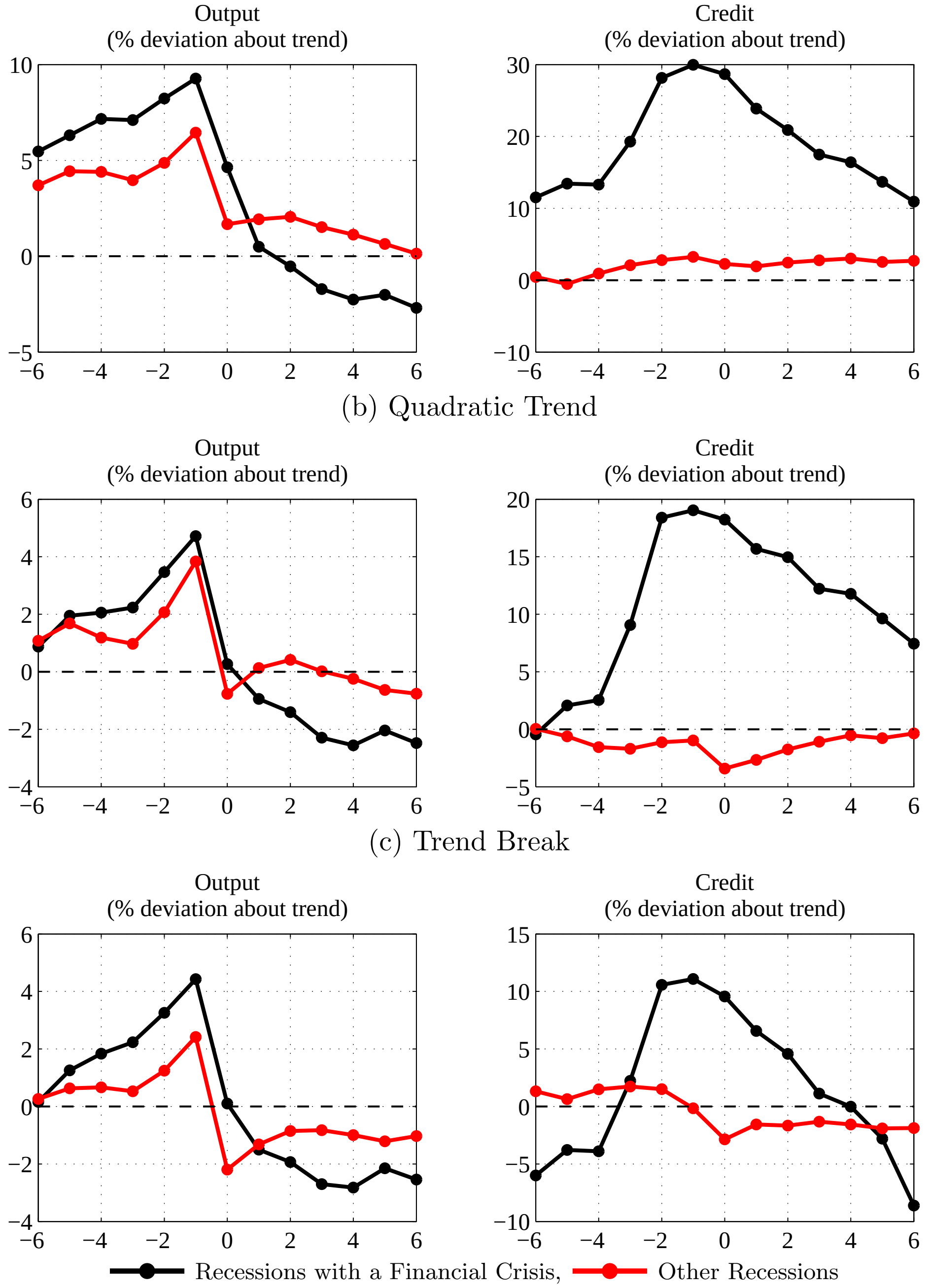


C Additional Tables and Figures

C.1 Recessions

Figure C.6: Financial versus normal recessions

(a) Linear Trend



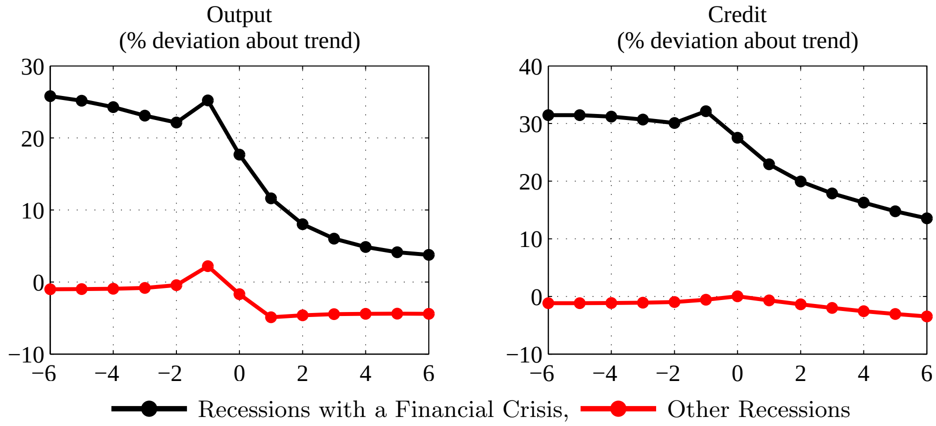
Note: The reported % deviations are the average % deviations from a deterministic trend (linear and quadratic) and from a trend before and after WWII. In this later case, we introduce dummies for the war years.

Table 6: Amplitude of Recessions

	Linear	Quadratic	Break
With SBC	11.96	7.27	7.25
Without SBC	6.31	4.61	4.61

Note: All numbers are reported in percentages from a deterministic trend (linear and quadratic), and from a trend before and after WWII. In this later case, we introduce dummies for the war years.

Figure C.7: Financial versus normal recessions (Model)



C.2 Dynamic Correlations

C.3 Bank Leverage

Figure C.9 reports the evolution of the interbank funding ratio in the US for the period 2004–2009. We consider several measures of the bank leverage, relating non core liabilities to various indicators (total equity, deposits, assets and core liabilities). Note that the closest measure to our model is given by the ratio of non core to core liabilities. Since we are mainly interested in the dynamics around the crisis, we normalized the level of each indicator to 100 in 2004Q1.

C.4 Benchmark Model vs RBC model: Typical Path

Figure C.8: Cross–correlogram GDP, Credit: Corr(y_t, c_{t-k}), $k = -4, 4$

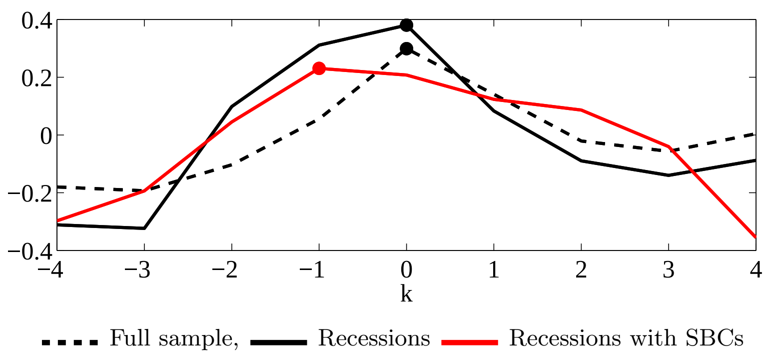


Figure C.9: Dynamics of Interbank Market Funfing Ratio (data)

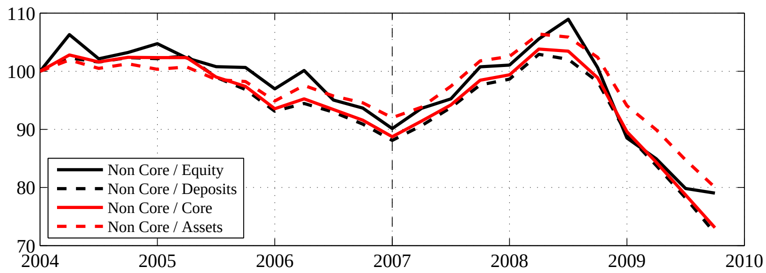
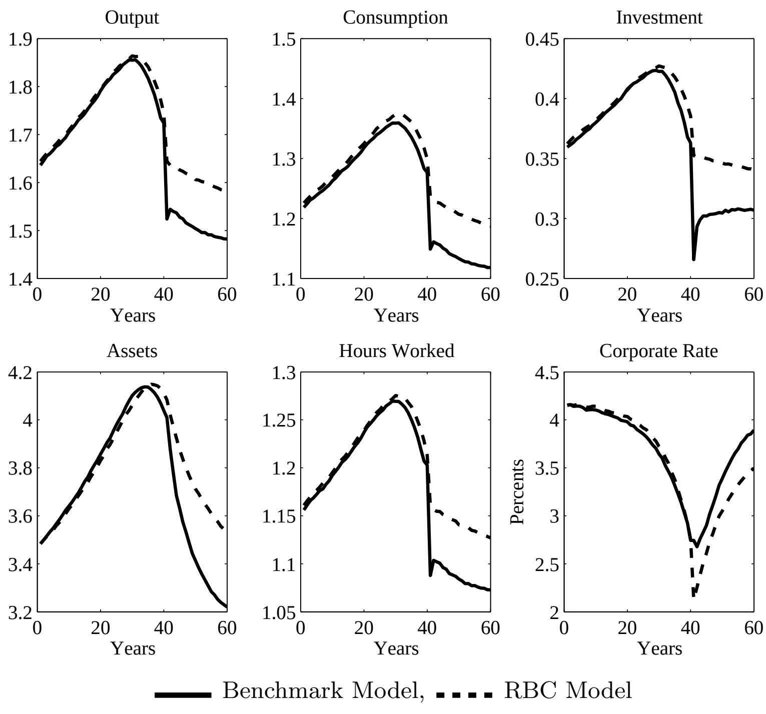


Figure C.10: Typical Path



C.5 Second Order Moments

Table 7 reports the second order moments for the main macroeconomic aggregates. Both the data and model simulations are HP–filtered with a coefficient $\lambda = 6.25$ as advocated by Uhlig and Ravn (2002).

Table 7: Second Order Moments

	US Data			Model		
	(a)	(b)	(c)	(a)	(b)	(c)
Output	2.17	1.00	–	2.70	1.00	–
Consumption	0.77	0.36	0.79	2.31	0.85	0.99
Investment	5.92	2.73	0.98	6.23	2.31	0.99
Hours	1.80	0.83	0.85	2.09	0.77	0.98

Note: (a): standard deviation, (b): relative standard deviation, (c): correlation with output.

Consumption is defined as as the sum of consumption of non durable goods and services per capita. Investment is the sum of gross fixed private investment and consumption of durable goods per capita. Hours are total hours in the non farm business sector divided by total civilian population. Output is the sum of consumption and investment.[47] We now report the exact construction of the data:

Consumption	=	`(PCNDA+PCESVA)/(CNP16OV*GDPA/GDPCA)`
Investment	=	`(GPDIA+PCDGA)/(CNP16OV*GDPA/GDPCA)`
Output	=	Consumption+Investment
Hours	=	`HOANBS/CNP16OV`

with

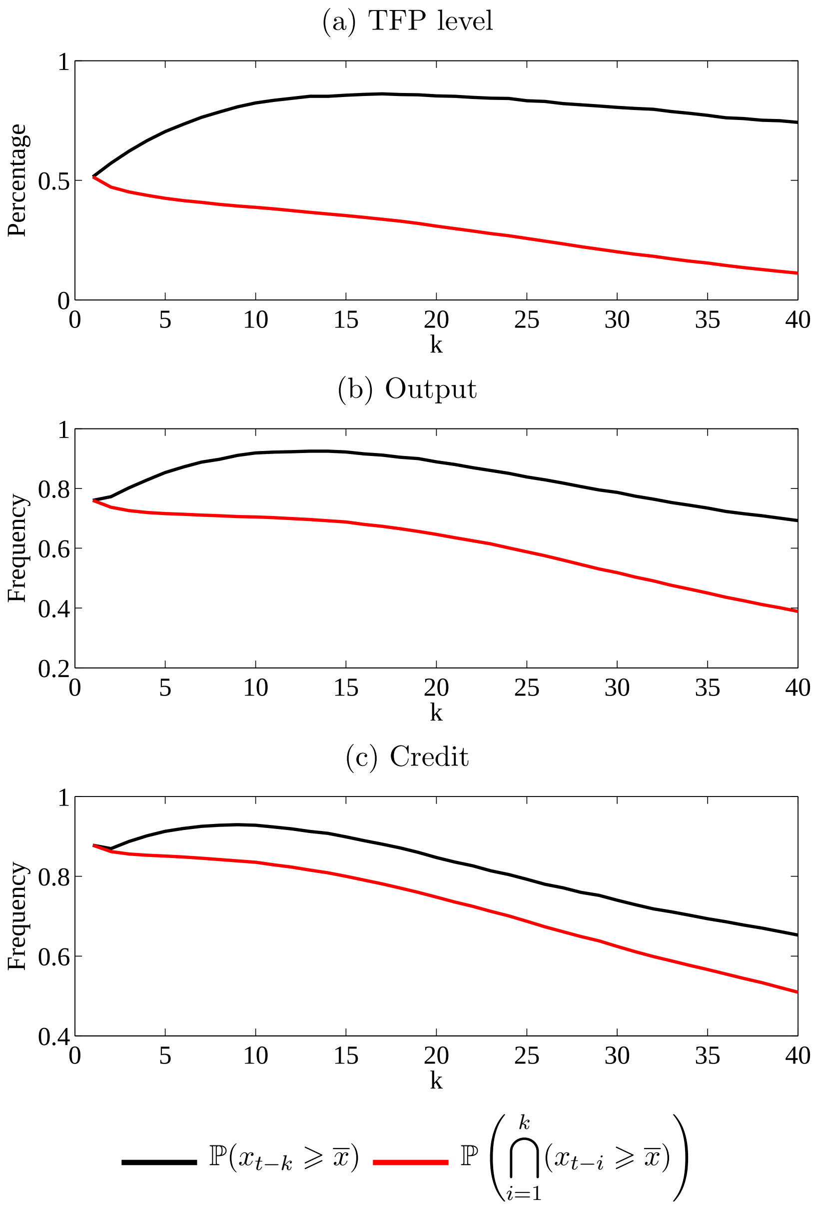
Mnemonic	Variable	Transformation
GDPA	Gross Domestic Product, Current Dollars	N/A
GDPCA	Gross domestic Product, Constant Dollars	N/A
PCESVA	Consumption of Services, Current Dollars	N/A
PCNDA	Consumption of non Durables, Current Dollars	N/A
PCDA	Consumption of Durables, Current Dollars	N/A
GPDIA	Gross Private Domestic Investment, Current Dollars	N/A
HOANBS	Total Hours Worked in Non Farm Business Sector	Average
CNP16OV	Civilian Population, 16 and over	End of Period

All data are downloaded from `http://research.stlouisfed.org/fred2/`

[47] Note that this is a limitation as it does not include the financial deadweigh loss of the model.

C.6 Crisis Break out During Credit Booms

Figure C.11: Percentage above average steady state before an SBC



Note: This figure reports the percentage of realizations of a variable above average steady state k periods before the crisis breaks out.

The Figure clearly reveals that SBCs are preceded by periods where total factor productivity, output, and credit are above the average steady state of the economy.

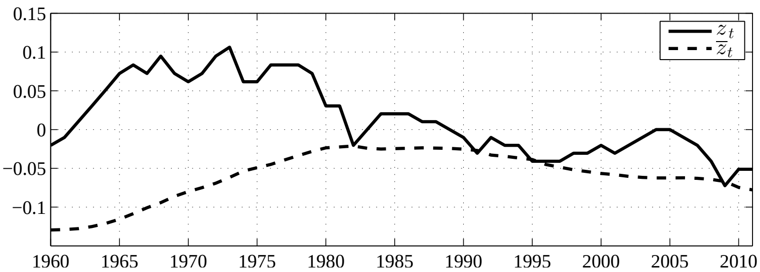
C.7 TFP and the Crisis

An alternative representation of the threshold leading to a crisis can be obtained relying on Total Factor Productivity. A crisis occurs when current TFP cannot "sustain" the level of assets in the economy. We can then define a productivity threshold, which, for our benchmark specification, is given by

$$\overline{z}_t = \Gamma_z a_t^{\frac{\nu(1-\alpha)}{1+\nu}} \text{ with } {}_z \equiv \left(\frac{\vartheta}{1-\alpha}\right)^{\frac{1-\alpha}{1+\nu}} \left(\frac{\overline{R}+\delta-1}{\alpha}\right)^{\frac{\alpha+\nu}{1+\nu}}$$

Figure C.12 reports the evolution of the level of TFP z_t and the threshold $\overline{z}_t$ from 1960 to 2011, and shows that the 1960s were a period in which, given the level of assets in the economy (and therefore $\overline{z}_t$), the productivity was high enough to sustain the level of assets in the economy. The credit boom of the 1960s was indeed a virtuous credit boom. Over time, the situation deteriorated as, with the increase in assets, the threshold productivity level rose, closing the gap between actual TFP and $\overline{z}_t$.

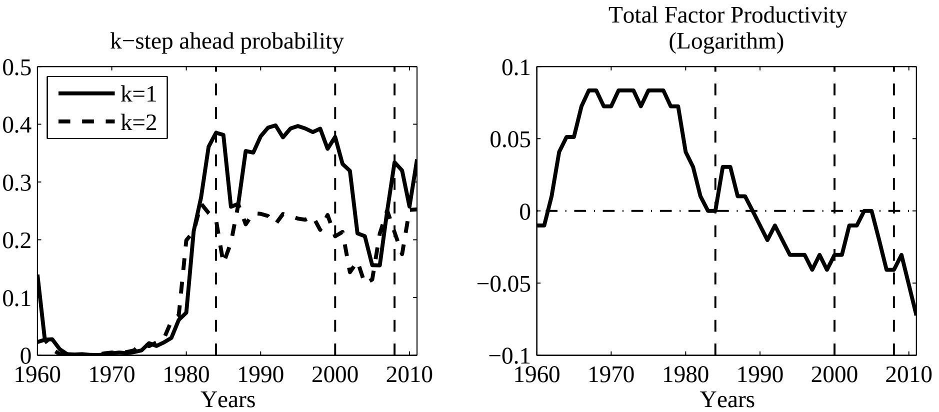
Figure C.12: TFP vs $\overline{z}_t$ in the US economy (1960–2011)



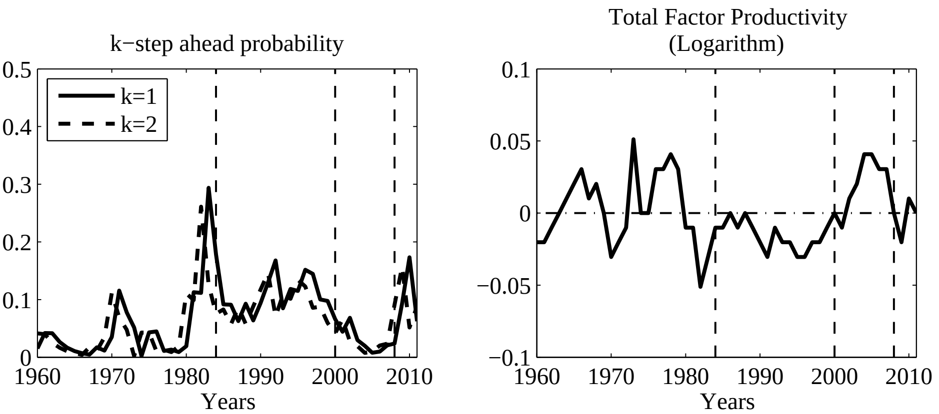
Note that in the main text we use the Solow residual as a measure of TFP. One may be concerned that this measure be contaminated by demand side effects. In order to correct for this potential problem, we report in panel (a) of Figure C.13 the probability we obtain using a measure of TFP corrected for utilization. As can be seen from the graphs, the results are essentially left unaffected by this correction. Some econometricians have argued that the evolution of TFP was affected by a structural break in its slope in 1973. Panel (b) of Figure C.13 reports the probability when TFP is corrected for such a break in the trend.[48] Although the evolution of probability differs in level from the previous measures, it still capture increases in the probability of occurrence of a crisis before it breaks out.

[48]Note however that this assumption is not consistent with the model in that the model assumes a constant (null) growth rate of TFP.

Figure C.13: k–step ahead Probabilities of a Financial Crisis (k=1,2)
(a) TFP corrected for utilization



(b) Break in 1973



Note: The vertical thin dashed lines correspond to the 1984 Savings & Loans, the 2000 dotcom and 2008 crises.

D The Model with Constant Saving Rate

D.1 Equations of the Model

1. $y_t = z_t k_t^{\alpha} h_t^{1-\alpha} + (\gamma + \delta - 1)(a_t - k_t)$
2. $R_t = \alpha k_t^{\frac{-\upsilon(1-\alpha)}{\upsilon+\alpha}} z_t^{\frac{1+\upsilon}{\upsilon+\alpha}} \left(\frac{1-\alpha}{\vartheta}\right)^{\frac{1-\alpha}{\upsilon+\alpha}} + 1 - \delta$
3. $h_t = \left(\frac{(1-\alpha) z_t}{\vartheta}\right)^{\frac{1}{\upsilon+\alpha}} k_t^{\frac{\alpha}{\upsilon+\alpha}}$
4. $\overline{a}_t \equiv ((1-\alpha)/\vartheta)^{\frac{1}{\upsilon}} \left(\alpha/\left(\overline{R}+\delta-1\right)\right)^{\frac{\upsilon+\alpha}{\upsilon(1-\alpha)}} z_t^{\frac{1+\upsilon}{\upsilon(1-\alpha)}}$
5. $a_{t+1} = i_t - (1-\delta) a_t$

If $a_t \leqslant \overline{a}_t$ (normal times)

- 6a. $k_t = a_t$
- 7a. $\frac{r_t}{R_t} = \int_{\overline{p}_t}^{1} p \frac{\mathrm{d}\mu(p)}{1-\mu(\overline{p}_t)}$
- 8a. $\overline{p}_t = \frac{\rho_t}{R_t}$
- 9a. $R_t = \frac{\rho_t}{\mu^{-1}\left(\frac{\rho_t-\gamma}{\rho_t-(1-\theta)\gamma}\right)}$, with $\rho_t > \overline{\rho}$
- 10a. $i_t = s\left(y_t - (R_t - r_t) a_t\right)$
- 11a. $y_t = c_t + i_t + (R_t - r_t) a_t$

If $a_t > \overline{a}_t$ (crisis times)

- 6b. $k_t = a_t - \mu(\gamma/R_t) a_t$
- 7b. $\frac{r_t}{R_t} = \frac{\gamma}{R_t}\mu(\gamma/R_t) + \int_{\gamma/R_t}^{1} p\,\mathrm{d}\mu(p)$
- 8b. $\overline{p}_t = n.a.$
- 9b. $\rho_t = \gamma$
- 10b. $i_t = s\left(y_t - (R_t - r_t) a_t + (R_t - \gamma)(a_t - k_t)\right)$
- 11b. $y_t = c_t + i_t + (R_t - r_t) a_t - (R_t - \gamma)(a_t - k_t)$

D.2 Implications for the Occurrence of Crises

Figure D.1: Decision Rules: Solow Model

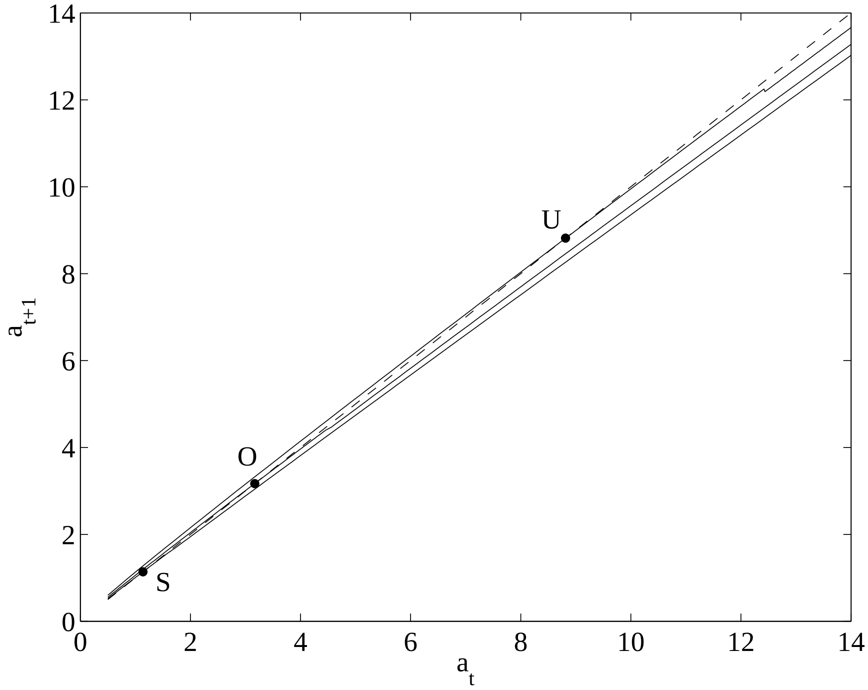
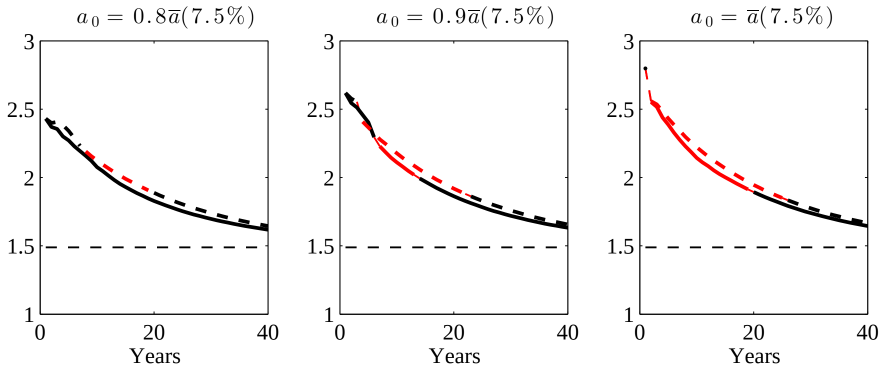


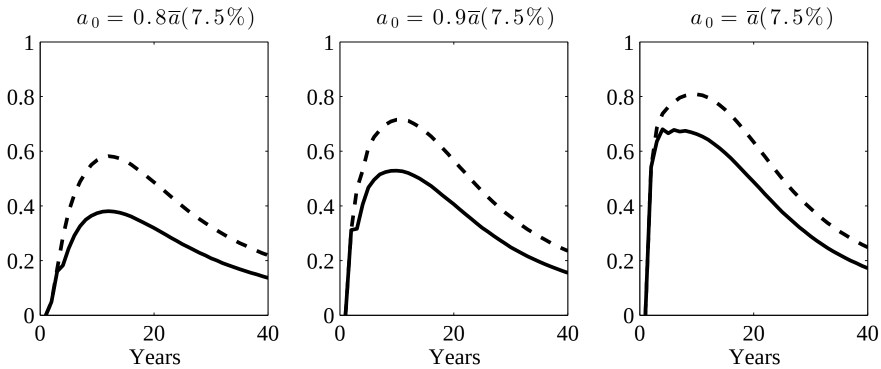
Figure D.2: The Role of Permanent Income Mechanisms

(a) Output



▬▬ (▪▪▪▪) Dynamics in normal times in the Solow version (Benchmark Model), ▬▬ (▪▪▪▪) Dynamics in a systemic banking crisis in the Solow version (Benchmark Model), - - - - long–run average. $\overline{a}(7.5\%)$ denotes the banks' absorptive when productivity is 7.5% above average.

Figure D.3: Frequency of SBCs: Sensitivity to Initial Conditions



▬▬ Solow Version, ▪▪▪▪ Benchmark Model.

Note: This figure reports the evolution of the frequency of SBCs during the transition toward the average steady state.

E An Alternative Model with Bank Leverage and Bank Defaults

In this section we endogenize bank leverage and study its dynamics along the typical path to crises. As in the text, we define bank deposits as the risk–free asset, and then assume a friction between the banks and the shareholder/household; this allows us to derive the optimal bank leverage. To introduce bank defaults, we further assume, unlike in the text, that a public deposit insurance scheme is in place that limits the banks shareholder liability, enables banks to issue more deposits, but also makes them prone to defaults.

Assume that a risk free asset exists in the economy, which we call "bank deposits". Let r^d_{t+1} be the risk–free (non–state contingent) gross return on deposits d_{t+1} and $r^e_{t+1}(p)$ be the *ex post* return on bank p's equity e_{t+1} at the end of period $t+1$ with, by definition, $a_{t+1} \equiv d_{t+1} + e_{t+1}$ and

$$r^e_{t+1}(p) \equiv \max\left(\frac{r_{t+1}(p)a_{t+1} - r^d_{t+1}d_{t+1}}{e_{t+1}}, 0\right), \tag{19}$$

where $r_{t+1}(p)$ is defined in (5). Whenever bank p's gross return on assets is smaller than the cost of deposits, bank p defaults. Its shareholder's limited liability then implies that bank p's return on equity is zero. Relation (19) also pins down the type of the marginal bank, $\widetilde{p}_{t+1}$, that defaults at the end of period t+1. Some banks default at the end of period $t+1$ only if the least efficient banks have a negative gross return on equity, *i.e.* only if $\rho_{t+1}a_{t+1} < r^d_{t+1}d_{t+1}$; hence, by definition,

$$\widetilde{p}_{t+1} \equiv r^{-1}_{t+1}\left(r^d_{t+1}\frac{d_{t+1}}{a_{t+1}}\right)\mathbb{I}_{\rho_{t+1}<r^d_{t+1}\frac{d_{t+1}}{a_{t+1}}}. \tag{20}$$

where $\mathbb{I}_{\rho_{t+1}<r^d_{t+1}\frac{d_{t+1}}{a_{t+1}}}$ is an indicator function that takes value one when $\rho_{t+1}a_{t+1} < r^d_{t+1}d_{t+1}$ and zero otherwise. The household's no–arbitrage condition between deposits and equity then requires that

$$\mathbb{E}_t(u'(c_{t+1})r^e_{t+1}) = r^d_{t+1}\mathbb{E}_t(u'(c_{t+1})), \tag{21}$$

which relation defines r^d_{t+1} in equilibrium and implies that $r^d_{t+1} < \mathbb{E}_t(r^e_{t+1})$. The deposit rate is determined at the end of period t. Finally, we define the *ex post* return of the household's asset portfolio, $\widetilde{r}_{t+1}$, as the sum of the returns on deposits and equity from all banks at the end of period t+1:

$$\widetilde{r}_{t+1} \equiv \int_0^1 \widetilde{r}_{t+1}(p)d\mu(p), \text{ where } \widetilde{r}_{t+1}(p) \equiv \frac{r^e_{t+1}(p)e_{t+1} + r^d_{t+1}d_{t+1}}{a_{t+1}} \geqslant r_{t+1}(p). \tag{22}$$

Importantly, the limited liability of the banks' shareholder (see (19)) implies that $\widetilde{r}_{t+1} \geqslant r_{t+1}$. And the household's saving decision is now given by

$$u'(c_t) = \beta\mathbb{E}_t(u'(c_{t+1})\widetilde{r}_{t+1}), \tag{23}$$

which suggests that the household tends to accumulate more assets than in the absence of limited liability (compare (23) with (4)). Now, because of the uncertainty surrounding the state of the economy in period $t+1$, there is a limit to the amount of deposits that banks can effectively issue. We assume that banks cannot issue more than a fraction κ of their total assets in the form of deposits, and thus face the following constraint:

$$d_{t+1}r^d_{t+1} \leqslant \kappa a_{t+1}, \tag{24}$$

where $\kappa > 0$. The above constraint is a feasibility constraint, which sets the maximum amount of deposits a bank can raise at the end of period t. If $\kappa \leqslant \gamma$, banks cannot commit themselves to paying back more than what the least efficient banks return in the worst state of the nature.[49] In this case,

[49] Given the data generating process of total factor productivity, the worst state of the nature is one where $\varepsilon_t \to -\infty$, $z_t \to 0 < \overline{z}_t$, and a banking crisis breaks out. In this case, the least efficient banks would store their assets and return γa_{t+1}. Such an event obviously comes with a zero probability but nevertheless belongs to the probability space. It follows that γa_{t+1} is the largest amount that a bank can commit itself to paying back with certainty. Put differently, this is also the largest amount of deposits a bank can issue in the absence of deposit insurance.

no bank ever defaults. To make things interesting we will assume that $\kappa > \gamma$, which means that banks may raise more deposits *ex ante* than they can repay in some states of the nature *ex post.* But in this case, for bank deposits to remain a risk–free asset (and to be able to write (21) in the first place) we assume that a deposit insurance scheme is in place that takes over the debts of the defaulting banks. For simplicity, we will further assume that deposit insurance is provided by a public sector that can credibly commit to guaranteeing the deposits by raising —if need be— lump–sum taxes from the household at the end of period $t+1$. In this context, $\kappa - \gamma$ must be interpreted as a measure of the scale of the deposit insurance provided by the government. Finally, we introduce a friction between the banks and the shareholder/household. We assume that banks are run by risk–neutral managers who choose the deposit to equity ratio so as to maximize their respective banks' discounted expected returns on equity:[50]

$$\max_{d_{t+1}/e_{t+1}} \beta \mathbb{E}_t(\widetilde{r}_{t+1}) + \beta(\mathbb{E}_t(\widetilde{r}_{t+1}) - r^d_{t+1})\frac{d_{t+1}}{e_{t+1}} \tag{25}$$

subject to feasibility constraint (24). Since $\mathbb{E}_t\widetilde{r}_{t+1} > r^d_{t+1}$ it is clear that managers want to raise as much deposits as possible and that, in equilibrium, constraint (24) binds. Hence,

$$\frac{d_{t+1}}{e_{t+1}} = \frac{\kappa}{\max(r^d_{t+1} - \kappa, 0^+)}. \tag{26}$$

Managers are not indifferent between equity and debt financing because —unlike what the shareholder would do— they do not discount the increased volatility of the return on equity that follows debt issuances. Given (26), one can now easily derive borrowing (lending) banks' optimal leverage ℓ^b_{t+1} (ℓ^l_{t+1}) as

$$\ell^b_{t+1} = \frac{d_{t+1}}{e_{t+1}} + \phi_{t+1}\left(1 + \frac{d_{t+1}}{e_{t+1}}\right) \text{ and } \ell^l_{t+1} = \frac{d_{t+1}}{e_{t+1}}. \tag{27}$$

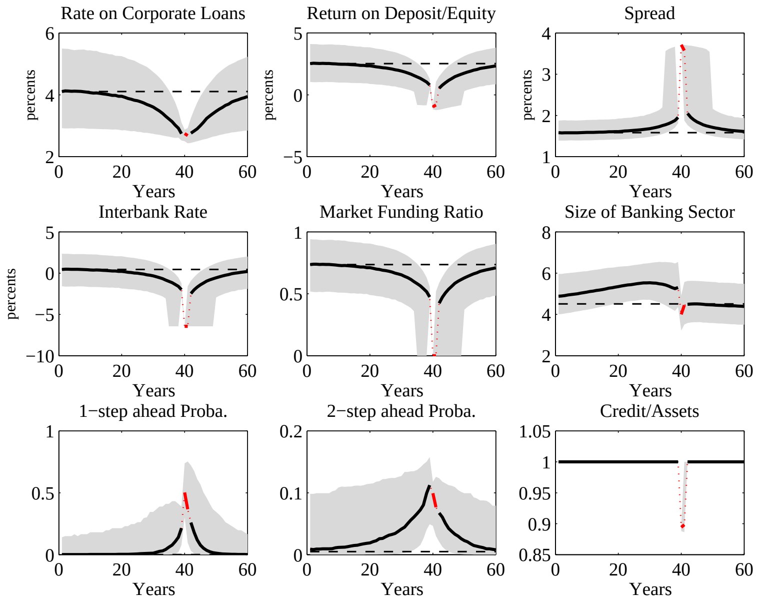
Denoting by Δ_{t+1} the mass of banks that default in the banking sector, one therefore gets:

$$\Delta_{t+1} = \mu\left(\widetilde{p}_{t+1}\right). \tag{28}$$

That the shareholder has limited liability changes the aggregate dynamics only marginally. The typical path to crisis is reported in Figures E.4 and E.5 for $\kappa = 0.94$.

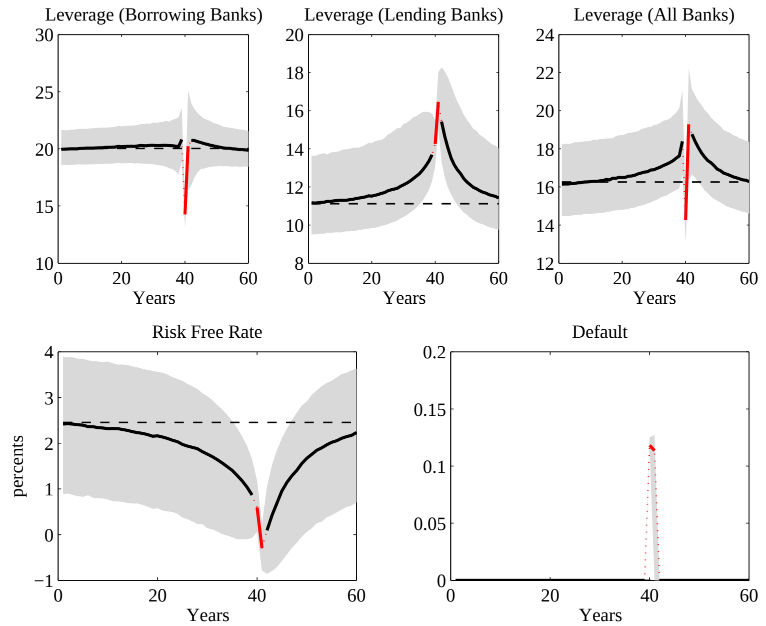
[50] More precisely, bank managers choose d_{t+1}/e_{t+1} to maximize $\mathbb{E}_t\beta r^e_{t+1}$ subject to the identities $\widetilde{r}_{t+1}a_{t+1} \equiv r^d_{t+1}d_{t+1} + r^e_{t+1}e_{t+1}$ and $a_{t+1} \equiv d_{t+1} + e_{t+1}$, with $r^e_t \equiv \int_0^1 r^e_t(p)\mathrm{d}\mu(p)$. If managers' incentives were aligned with the banks' shareholders, then managers would instead maximize $\mathbb{E}_t(\beta\frac{u'(c_{t+1})}{u'(c_t)}r^e_{t+1})$ and be indifferent between equity and deposits.

Figure E.4: Typical path (I)



Dynamics in normal times, Dynamics in a systemic banking crisis, long–run average.

Figure E.5: Typical path (II)



Dynamics in normal times, Dynamics in a systemic banking crisis, long–run average.

F The Model with Direct Finance

In this section, we consider a version of the model in which firms have partial access to direct financing. More precisely, we assume that households deposit a fraction $1-\omega \in (0,1)$ of their assets in banks, and can use the remaining fraction ω to directly finance firms' projects. For the sake of simplicity, we ignore labor supply decisions. This leads to a slight modification in the household's budget constraint, which now reads

$$a_{t+1} + c_t = \omega R_t a_t + (1-\omega) r_t a_t + w_t + \pi_t$$

implying that the Euler equation now reads

$$u'(c_t) = \beta \mathbb{E}_t \left[u'(c_{t+1}) \left(\omega R_{t+1} + (1-\omega) r_{t+1} \right) \right].$$

This implies that the banking sector only has access to a fraction of assets. The general equilibrium of this economy then reads

1. $y_t = z_t k_t^{\alpha} + (\gamma + \delta - 1)(a_t - k_t)$
2. $R_t = \alpha k_t^{\alpha - 1} + 1 - \delta$
3. $c_t^{-\sigma} = \beta \mathbb{E}_t \left[c_{t+1}^{-\sigma} (\omega R_{t+1} + (1-\omega) r_{t+1} \right]$
4. $\overline{a}_t \equiv \left(\alpha z_t / \left(\overline{R} + \delta - 1 \right) \right)^{\frac{1}{1-\alpha}}$
5. $i_t = a_{t+1} - (1-\delta) a_t$

If $a_t \leqslant \overline{a}_t$ (normal times)

6*a*. $k_t = a_t$

7*a*. $\dfrac{r_t}{R_t} = \displaystyle\int_{\overline{p}_t}^{1} p \frac{\mathrm{d}\mu(p)}{1 - \mu(\overline{p}_t)}$

8*a*. $\overline{p}_t = \dfrac{\rho_t}{R_t}$

9*a*. $R_t = \dfrac{\rho_t}{\mu^{-1}\left(\frac{\rho_t - \gamma}{\rho_t - (1-\theta)\gamma} \right)}$, with $\rho_t > \overline{\rho}$

10*a*. $y_t = c_t + i_t + (R_t - r_t)(1-\omega) a_t$

If $a_t > \overline{a}_t$ (crisis times)

6*b*. $k_t = a_t - (1-\omega)\mu(\gamma / R_t) a_t$

7*b*. $\dfrac{r_t}{R_t} = \dfrac{\gamma}{R_t} \mu(\gamma / R_t) + \displaystyle\int_{\gamma / R_t}^{1} p \, \mathrm{d}\mu(p)$

8*b*. $\overline{p}_t = n.a.$

9*b*. $\rho_t = \gamma$

10*b*. $y_t = c_t + i_t + (R_t - r_t)(1-\omega) s_t + (R_t - \gamma)(s_t - k_t)$

Given that this version of the model does not include endogenous labor supply, the parameters of the model have to be adjusted. We recalibrate the model using the same approach as the one we used in the body text, assuming, for our benchmark calibration, that agents have no access to direct finance ($\omega = 0$). The parameters are reported in Table 8. Figures F.6–F.7 report the change in long–run average across 500,000 simulations of the model as we vary the share of direct finance — *i.e.* increase ω. First and foremost, the level of assets, and therefore credit, increases as the share of direct finance increases. On the one hand, increase in direct finance limits the degree of financial frictions and systemic risk in the economy. This effect plays against the increase in the level of accumulation, as this should weaken savings for precautionary motives. On the other hand, the average return to asset accumulation increases for the household with ω. This can be seen from the Euler equation in which the return to accumulation is given by

$$\omega R_{t+1} + (1-\omega) r_{t+1} \geqslant r_{t+1}$$

Table 8: Calibration

Parameter	Values
Discount factor	$\beta = 1/1.03$
Risk aversion	$\sigma = 8.00$
Capital elasticity	$\alpha = 0.300$
Capital depreciation rate	$\delta = 0.100$
Standard dev. productivity shock	$\sigma_z = 0.018$
Persistence of productivity shock	$\rho_z = 0.900$
Bank distribution - $\mu(p) = p^{\lambda}$	$\lambda = 30$
Diversion cost	$\theta= 0.1$
Storage technology	$\gamma = 0.955$

Hence compared to a situation without direct finance, intertemporal substitution motives leads to an increase in asset accumulation. Furthermore, the larger returns to accumulation exert a positive wealth effect which further magnifies accumulation. Since the absorptive capacity of the banking sector is not affected by the degree of direct finance, the gap between assets and absorptive capacity closes, making it more likely for the economy to experience a crisis. Therefore the probability of as SBC increases with ω (see left panel in Figure F.7). However, the median amplitude of the crisis diminishes. (Notice that in the limit case $\omega = 1$ interbank market freezes become immaterial since there is no financial intermediation.) As asset accumulation and credit increase, the rate on corporate loans decreases. Access to direct finance limits the need for the interbank market as less demand for funds is addressed to banks. The interbank rate decreases.

Figure F.6: Variation in Direct Finance I

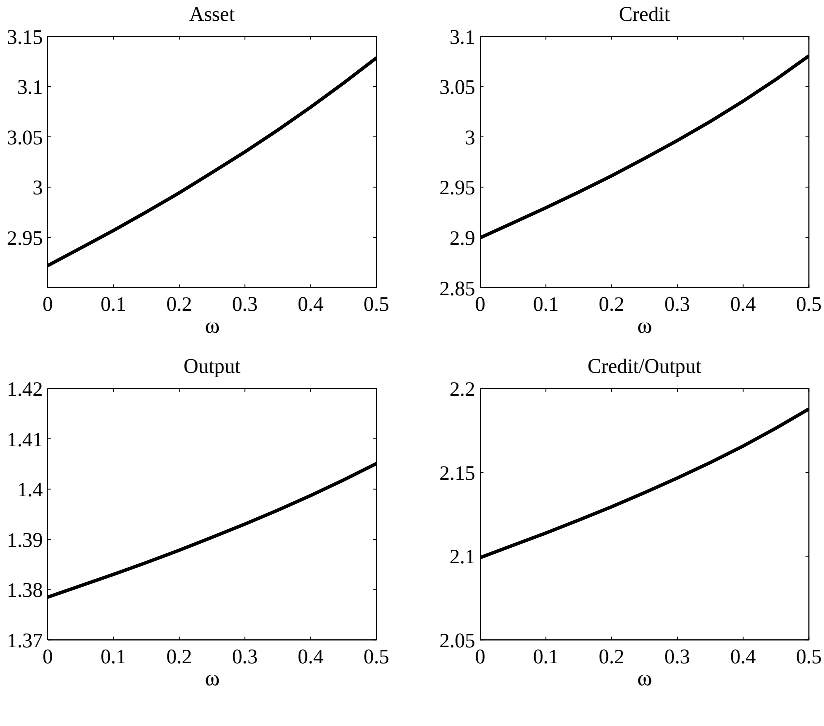


Figure F.7: Variation in Direct Finance II

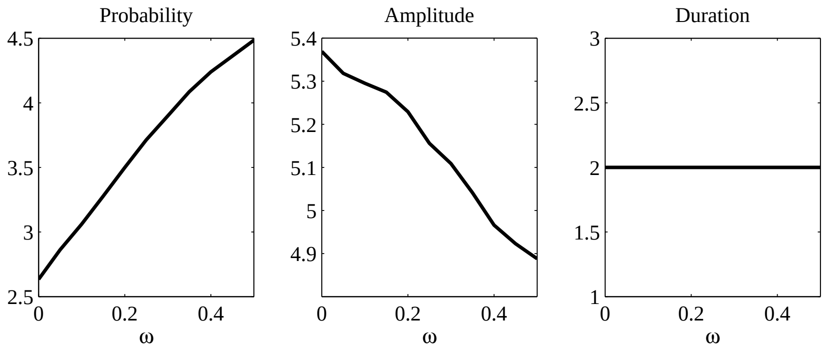


Figure F.8: Variation in Direct Finance III

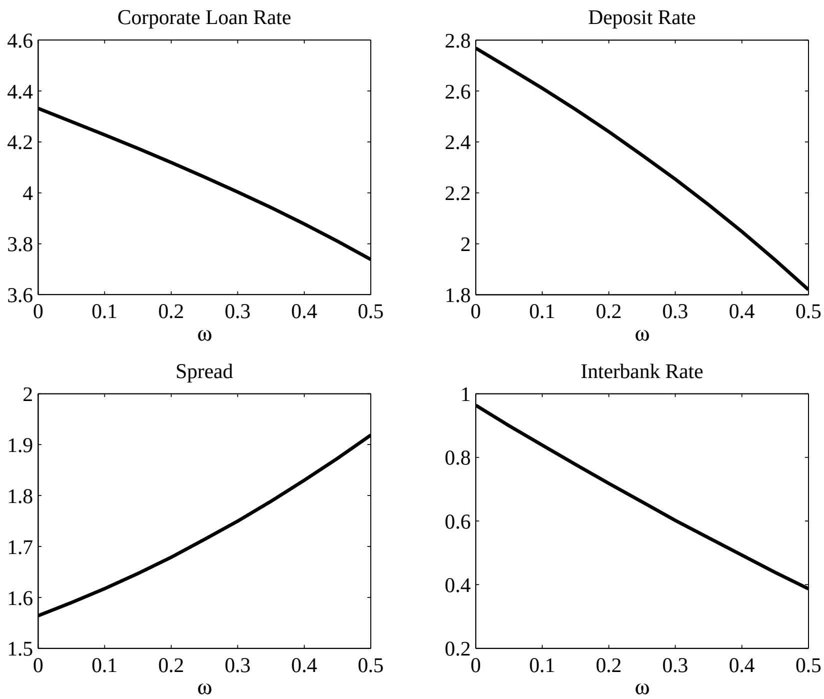


Table 9: Variations in Direct Finance

ω	0.00	0.25	0.50
Assets (s_t)	2.92	3.01	3.13
Credit (k_t)	2.90	2.98	3.08
Output (y_t)	1.38	1.39	1.41
k_t/y_t	2.10	2.14	2.19
Probability	2.63	3.71	4.49
Amplitude	5.37	5.16	4.89
Duration	2.00	2.00	2.00
Interbank Rate (ρ_t)	0.96	0.66	0.39
Corporate Rate (R_t)	4.33	4.06	3.74
Deposit Rate (r_t)	2.77	2.35	1.82
Spread	1.56	1.71	1.92

Note: This table reports averages across 500,000 simulations, except for amplitude and durations which correspond to the median of the distribution.