Financial and Regulatory Cycles

Frédéric Boissay* Fabrice Collard† Ingo Fender‡

December 31, 2018

Very preliminary and incomplete
Do not quote

Abstract

We study the endogenous dynamics of savings, bank leverage, and shadow banking in the run–up to financial crises, through the lens of a dynamic macro-model. We classify endogenous financial crises on the basis of the type of imbalances that lead to them. Some crises follow the development of a savings glut ("push" factor); some follow a banking glut, characterised by regulatory arbitrage and shadow banking activities ("pull" factor). The simulations of the model calibrated on the US indicate that one third of banking crises are due to regulatory arbitrage and one half to excess savings, while the rest are due to concomitant savings and banking gluts. Those imbalances have similar effects —a deep recesion— but call for varied policy interventions.

Keywords: xxx

JEL Class.: E60 – G18 – G28.

*Bank for International Settlements, Centralbahnplatz 2, 4055, Basel, Switzerland. Email: frederic.boissay@bis.org.
†Toulouse School of Economics. Email: fabrice.collard@gmail.com. URL: http://fabcol.free.fr.
‡Bank for International Settlements, Centralbahnplatz 2, 4055, Basel, Switzerland. Email: ingo.fender@bis.org.

1 Introduction

The 2008 US financial crisis is usually ascribed to one of two imbalances: a saving glut or a banking glut.

In the saving glut story (Bernanke, 2005), banks are seen as channeling savings to investments through their retail lending activity. Imbalances arise when there is too much savings relative to the economy's investment capacity. Excess savings weigh on lending rates and crowd in lower quality (e.g. subprime) borrowers, which reduces aggregate productivity and leaves the economy more vulnerable to adverse shocks. The saving glut story is one of excess *net* capital flowing *into* the financial sector.

The banking glut story (Shin, 2011), in contrast, is one of excess *gross* capital flows *within* the financial sector. Imbalances arise when banks take advantage of loopholes in the regulation. By securitizing and selling loans off their balance sheet to non–regulated financial entities ("shadow" banks), banks economize on regulatory capital, and free up lending capacity.

Both stories feature a credit boom and have the same consequences: a credit crunch and a severe recession. But the reason behind the boom is different. In the saving glut story, it is excess savings due —*inter alia*— due to aging population, a current account deficit, or excess demand for safe assets in emerging market economies. In this case, the culprit is to be found outside the financial sector. In the banking glut story, the credit boom is due to regulatory arbitrage by the financial sector.

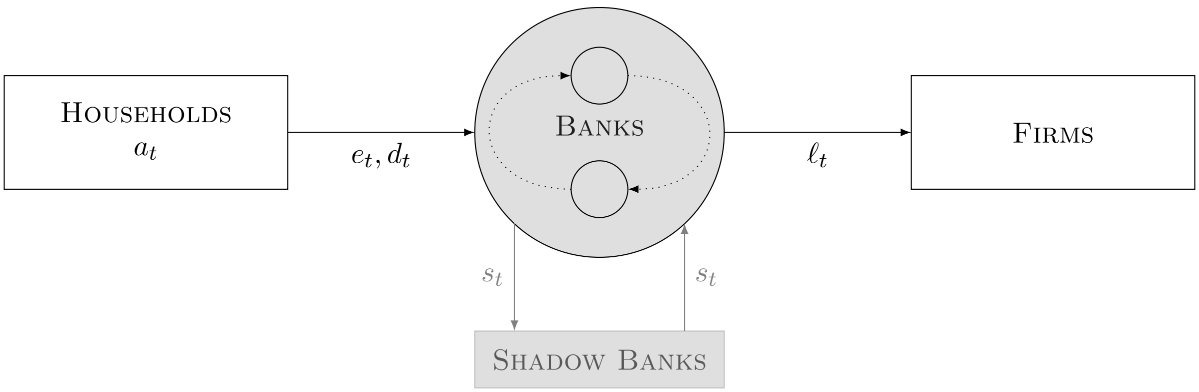
We study the endogenous dynamics toward banking crises through the lens of a dynamic general equilibrium model. In the model, banking crises take the form of a sudden wholesale funding liquidity stress for banks. They may break out due to excess savings and/or regulatory arbitrage. We use the model to discuss the effects of bank capital regulation, depending on the type of financial imbalances. Tighter capital requirements helps to cope with a savings glut. But too tight a regulation may also give banks incentive to get around it, by creating a shadow banking sector. Implicit in our discussion is indeed the assumption that a rapid pace of financial innovations enables financial institutions to create and exploit regulatory arbitrage opportunities.

To be completed...

2 Model

We consider an economy populated with mass one continuums of atomistic risk averse households, firms, and banks. At the end of period $t-1$, households consume, c_{t-1}, and save a_t. They may invest a_t in bank deposits, d_t, and/or bank equity, e_t. Banks, in turn, invest into corporate loans, ℓ_t. Under some circumstances that we describe later, banks may also invest in securities s_t issued by other banks as well as in securities issued by "shadow banks".[1] Figure 1 summarizes the financial flows between the agents in the economy.

Figure 1: Financial Flows



2.1 The Firm

There is a representative firm, established at the beginning of period t and closed at the end of period t. The firm produces $z_t k_t^\alpha$ homogeneous goods with k_t units of capital. Productivity, z_t, is stochastic and follows the process

$$\log(z_t) = \rho_z \log(z_{t-1}) + \epsilon_t^z, \tag{1}$$

where the aggregate productivity shock, ϵ_t^z, is log–normally distributed, and drawn at the beginning of period t (see Figure 5). The firm acquires its capital at the beginning of period t, before the realization of the productivity shock. The firm finances its capital by borrowing ℓ_t from the banks, at rate r_t^ℓ. It maximises its discounted expected profit,

$$\max_{k_t, \ell_t} \mathbb{E}_{t-1} \left(\Psi_{t-1,t} \pi_t \right), \tag{2}$$

where $\pi_t \equiv z_t k_t^\alpha + (1-\delta) k_t - r_t^\ell \ell_t$, given the depreciation rate δ and under the constraint $k_t \leq \ell_t$. The term $\Psi_{t-1,t} \equiv \beta \frac{u'(c_t)}{u'(c_{t-1})}$ is the household's discount factor (see Section 2.2 below). $\mathbb{E}_{t-1}(\cdot)$ denotes the expectation operator taken over the aggregate shocks, conditional on the information available at the end of period $t-1$. The first order conditions yield:

$$k_t = \ell_t \tag{3}$$

$$r_t^\ell = \alpha z_t k_t^{\alpha-1} + 1 - \delta. \tag{4}$$

The interest rate on loans, r_t^ℓ, is assumed to be state contingent, and the profit, π_t, is assumed to be rebated lump–sum to households.

[1] Shadow banks in our model are akin to what is usually referred to as special purpose vehicles.

2.2 Households

There is a mass one continuum of atomistic households. Households take their decisions in two stages. First, at the end of period $t-1$, they consume c_{t-1} and save a_t. Second, at the beginning of period t, they decide whether to invest a_t in bank deposits, d_t, or in bank equity, e_t, with $a_t = d_t + e_t$. Households take their portfolio allocation decisions after they learn of their financial transaction costs. Indeed, we assume that investing in bank deposits or equity is costly, and that the cost is idiosyncratic. We refer to household $(q^d, q^e) \in [0,1]^2$ as the household with transaction cost $1-q^d$ on deposits and $1-q^e$ on bank equity. For household (q^d, q^e), the net unit return on deposits is $q^d r_t^d$, and that on equity is $q^e r_t^e$ —where r_t^d and r_t^e refer to gross returns. To fix ideas, we associate those costs with the effort of screening, the cost of financial intelligence, financial advisor fees, or any other cost imposed by resolving contract enforcement problems with banks.[2] Given that returns are linear, a household specializes and invests a_t in either deposits or equity. That is, each household has her own "preferred–habitat" asset.[3] The decision to invest into deposits is denoted by $\mathbb{1}_t^d$, with $\mathbb{1}_t^d = 1$ if the household invests and $\mathbb{1}_t^d = 0$ otherwise. Similarly, the decision to invest into equity is denoted by $\mathbb{1}_t^e$, with $\mathbb{1}_t^e = 1 - \mathbb{1}_t^d$. The household takes her decisions so as to maximize her expected utility:[4]

$$\max_{\{a_{t+j}, c_{t+j-1}\}_{j=0,\ldots,\infty}} \mathbb{E}_{t-1} \sum_{j=0}^{j=\infty} \beta^j \left[\max_{\{\mathbb{1}_{t+j}^d\}_{j=0,\ldots,\infty}} u(c_{t+j-1}) \right] \tag{5}$$

subject to the constraints:

$$c_{t-1} + a_t = r_{t-1} a_{t-1} + \pi_{t-1}, \tag{6}$$

$$r_t a_t \equiv \mathbb{Q}_t^d r_t^d d_t + \mathbb{Q}_t^e r_t^e e_t, \tag{7}$$

$$d_t = a_t \int_0^1 \int_0^1 \mathbb{1}_t^d \mathrm{d}\mu_d(q^d) \mathrm{d}\mu_e(q^e), \tag{8}$$

$$e_t = a_t - d_t, \tag{9}$$

where $u(\cdot)$ satisfies the usual regularity conditions; $\mathbb{Q}_t^d r_t^d d_t$ and $\mathbb{Q}_t^e r_t^e e_t$ are the average net returns on deposits and equity, i.e. net of the transaction costs. The maximization problem is solved backward, starting with the second stage.

[2] Of course, those costs may also be affected by the institutional environment, like the quality of the legal system or the intensity of bank supervision. Our model does not capture those. More generally, we assume away any "residual" financial frictions between households and banks, which will allow us to focus on frictions within the banking sector (see Section 2.4). Financial transaction costs are relevant in practice. **?** for example report that 73% of the US households consult a financial advisor before purchasing shares or mutual funds. In the US mutual fund industry, the so–called "expense ratio 12b-1" is typically between 0.25% and 1% (the maximum allowed) of a fund's assets under management. **?** report that, including all management fees, advised portfolios cost Canadian households 2.5% of their assets under management per year. Formally, q^d and q^e are drawn randomly and independently from cumulative distributions $\mu_d(q^d)$ and $\mu_e(q^e)$, respectively.

[3] Note that bank capital regulation will therefore be costly for depositors.

[4] For simplicity, we assume that households pool their resources at the end of the period, and therefore all consume the same amount of final goods. In other terms, households perfectly insure each other against the idiosyncratic transaction costs.

Second Stage: Asset Portfolio Choice Given her transaction costs, household (q^d, q^e) deposits a_t with the banks if and only if her expected net return on deposits is higher than that on equity:[5]

$$\mathbb{1}_t^d = 1 \Leftrightarrow q^e < q^d \frac{\mathbb{E}_{t-1}\left(\Psi_{t-1,t}\right) r_t^d}{\mathbb{E}_{t-1}\left(\Psi_{t-1,t} r_t^e\right)}. \tag{10}$$

Based on (10), the total amount of wealth invested into deposits is

$$d_t = a_t \int_0^1 \mu_e \left(\min \left[1, q^d \frac{\mathbb{E}_{t-1}\left(\Psi_{t-1,t}\right) r_t^d}{\mathbb{E}_{t-1}\left(\Psi_{t-1,t} r_t^e\right)} \right] \right) \mathrm{d}\mu_d(q^d), \tag{11}$$

the remaining wealth being invested into equity ($e_t = a_t - d_t$). The average transaction costs associated with bank deposits and equity are given by (one minus):

$$\mathbb{Q}_t^d \equiv \frac{a_t}{d_t} \int_0^1 q^d \mu_e \left(\min \left[1, q^d \frac{\mathbb{E}_{t-1}\left(\Psi_{t-1,t}\right) r_t^d}{\mathbb{E}_{t-1}\left(\Psi_{t-1,t} r_t^e\right)} \right] \right) \mathrm{d}\mu_d(q^d) \tag{12}$$

and

$$\mathbb{Q}_t^e \equiv \frac{a_t}{e_t} \int_0^1 q^e \mu_d \left(\min \left[1, q^e \frac{\mathbb{E}_{t-1}\left(\Psi_{t-1,t} r_t^e\right)}{\mathbb{E}_{t-1}\left(\Psi_{t-1,t}\right) r_t^d} \right] \right) \mathrm{d}\mu_e(q^e). \tag{13}$$

First Stage: Consumption/Savings The household chooses the levels of consumption/savings, anticipating the second stage portfolio choices. This yields

$$\mathbb{E}_{t-1}\left(\Psi_{t-1,t} r_t\right) = 1 \tag{14}$$

as optimality condition.

2.3 Shadow Banks

In the model, "shadow" banks are unregulated entities that can neither service corporate loans nor extract cash from the firm. Those entities are created endogenously and *ex nihilo*, as banks' response to regulation. Shadow banks purchase loans from banks *before* banks are being supervised, against securities that return r_t^s. And they resell the loans to banks *after* banks have been supervised (see Figure 5), against claims that return r_t^i. They are nothing but "special purpose vehicles" that permit banks to window–dress their balance sheets for the supervisory review and circumvent regulation (see next section).[6] The objective of a shadow bank is to maximise its expected net return with respect to the volume of securities it issues, s_t:

$$\max_{s_t} \mathbb{E}_{t-1} \left(\Psi_{t-1,t} (r_t^i - r_t^s) \right) s_t. \tag{15}$$

[5]This condition takes into consideration that r_t^d is pre–determined at the end of period $t-1$, at the time deposits are made.

[6]Although shadow banks may resemble the special purpose vehicles that were created in the run–up to the Great Financial Crisis, we shall nonetheless not pretend to give a thorough account of the disparate shadow banking sector here. For example, shadow banks in our model merely repackage and resell the loans to banks. Yet, in practice, securitisations are usually tranched, with the tranches then potentially being sold to different types of investors. Before the Great Financial Crisis, a lot of these tranches were actually sold to other banks. We abstract from tranching here to focus primarily on the regulatory arbitrage motive of securitisation.

We assume that shadow banks break even *ex post*:

$$r_t^s = r_t^i. \tag{16}$$

2.4 Banks

The banking sector is composed of a continuum of one–period lived banks with total mass one.[7] At the beginning of the period, banks are identical. They take their decisions in two stages. In the first stage, the representative bank raises d_t deposits and lends $d_t + e_t$ to the firm, taking its equity funding, e_t, as given. The bank may also sell s_t loans to shadow banks and keep x_t loans on its balance sheet, with

$$x_t = d_t + e_t - s_t. \tag{17}$$

In exchange of the loans, the bank receives a security issued by the shadow banks that returns r_t^s at the end of the period. The representative bank may in this way "securitize" part of its loan book, though we assume that there is a limit on the fraction of the loan book that can be securitized,

$$s_t \leq \theta(d_t + e_t). \tag{18}$$

In the above constraint, parameter θ can be seen as a technological limit to securitization.

In a second stage, banks draw random idiosyncratic loan servicing costs. Indeed, we assume that banks must service the loans (or monitor the firms) throughout the period, and that the monitoring cost varies randomly across banks. Depending on the bank, the unit return on corporate loans net of this cost ranges from γ to r_t^ℓ.[8]

Assumption 1 ***(Banks' loan servicing costs)*** *At the beginning of period t, each bank learns of its return on loans net of its idiosyncratic loan servicing cost, $q^\ell r_t^\ell$.*

For the banks that do not incur any servicing cost, the net return is equal to the gross return, r_t^ℓ. For those that incur the highest servicing cost, the net return is γ.[9] For the exposition, we will express the net return as a fraction q^ℓ of the gross return, with $q^\ell \in \left[\frac{\gamma}{r_t^\ell}, 1\right]$, and refer to bank q^ℓ as the bank with net return $q^\ell r_t^\ell$. The q^ℓs are drawn from a cumulative distribution $\mu_\ell(q^\ell)$. A high–q^ℓ bank is thus more efficient in servicing loans (or monitoring firms) compared to a low–q^ℓ bank. The q^ℓs are drawn at the same time as the aggregate shocks.

[7]Note that our model with short–lived banks is isomorphic to one, where banks live infinitely. The reason is that, in our model, banks have a frictionless access to the equity market and no reputation risk (indeed, they draw their types afresh at the beginning of every period). Given that each household is risk averse and can freely reallocate equity and deposits across banks, it is optimal for it to diversify its equity holdings and deposits across the continuum of banks every period. The upshot is that, even if they lived infinitely, banks would start out every period with identical balance sheets, like new-born banks. In other terms, bank equity is not a state variable. In this respect, our approach is similar to **?** or **?**, but differs from other models with infinitely–lived banks that assume the absence of equity markets, like **?**. In this latter model, banks accumulate equity by retaining earnings. We do not make this assumption here, because this would bar us from studying the effects of bank capital regulation.

[8]Inada conditions on firms' technology will ensure that, in equilibrium, $r_t^\ell > \gamma$.

[9]Implicit is the assumption that a firm must be monitored, for its production to be completed.

When a bank learns of its q^ℓ, it already made the loans, and has two options going forward. One option is to keep and service its loans —or even purchase loans from other banks and service those loans as well. This will typically be the case if the bank has a high q^ℓ. The other option is to sell its loans to other banks. This will be the case if the bank has a low q^ℓ. Bank heterogeneity thus gives rise to a secondary market for corporate loans, where low–q^ℓ banks sell their loans to high–q^ℓ banks, against the promise to receive r_t^i as unit gross return. Alternatively, one can view this market as a wholesale funding market, or as an interbank market, where high–q^ℓ banks "borrow" from low–q^ℓ banks. In what follows, we will interchangeably refer to the seller of a loan on the secondary market as a "lender" on the interbank market, and to the buyer as a "borrower".

Clearly, in a frictionless world, only the banks with $q^\ell = 1$ would borrow and service corporate loans. Those banks would purchase loans from the rest of the banking sector, and —given their small size— would be infinitely leveraged. To motivate banks' wholesale funding limit, we follow **?**, and introduce the following friction. First, we assume that its q^ℓ is known to bank q^ℓ only, and therefore that information is asymmetric. Second, we assume that a bank has the possibility to extract up to γ of non–seizable cash from the firm, per unit of loan. That is, a bank that absconds gets as much as net return as the lowest q^ℓ bank; for the latter, the opportunity cost of defaulting is zero.

Assumption 2 ***(Agency problem)*** *The idiosyncratic loan servicing costs are private information. A bank has the possibility to extract up to γ of cash per unit of loan. This cash is not traceable/seizable by creditors.*

Because the cash extracted cannot be seized by other banks, the bank may renege on its debt, and abscond. As a result, γ influences the ease with which banks can raise funding on wholesale funding markets, *i.e.* banks' "funding liquidity". Let $\mathbb{1}_t^i$ be the decision to purchase loans on the secondary market, and ϕ_t the amount of loans purchased, with $\phi_t \geq 0$. Sellers/lenders protect themselves against a default by taking care never allow ϕ_t to exceed a certain threshold. Since the q^ℓs are not publicly observable, they must in addition make sure that even the banks with the lowest q^ℓ do not purchase loans to later abscond. Indeed, such banks should be selling —not purchasing— loans. A bank with the lowest q^ℓ, for example, earns $r_t^i x_t + r_t^s s_t - r_t^d d_t$ as net pay–off if it sells its x_t loans on the secondary market. If instead this bank purchases ϕ_t loans, extracts the non–seizable cash, and defaults strategically on its interbank loan, it earns $\gamma(x_t + \phi_t) - \kappa_t$, where κ_t is a time–varying default cost.[10] We also assume that deposits are defaultable and that creditors can size the cash flow from securities, $r_t^s s_t$, at no cost.[11] Given all this, the incentive constraint that

[10] This cost is admittedly *ad hoc*. It is not a fundamental element of the model, and it does not affect banks' marginal decisions. We only introduce this cost to gain some flexibility for the calibration of the model. In relation (34), we model it as an iceberg —and therefore time–varying— cost.

[11] What is important for capital regulation to have an effect is that *some* —not necessarily *all*— deposits be defaultable.

makes sure that no bank defaults is:

$$\gamma(x_t + \phi_t) - \kappa_t \leq r_t^i x_t + r_t^s s_t - r_t^d d_t. \tag{19}$$

The bank's opportunity cost of absconding, on the right hand side, goes up with the interbank rate, r_t^i, meaning that any rise in r_t^i enhances the bank's funding liquidity. As we explain later, the fact that the bank does not internalize this effect will be the reason why banks must be regulated. We assume that the regulator can only constrain banks' balance sheets as follows.

Assumption 3 ***(Regulation)*** *Banks must hold at least τ equity per unit of corporate loan at the beginning of every period:*

$$\frac{e_t}{x_t} \geq \tau. \tag{20}$$

The regulator makes sure that banks comply with regulation by supervising them. However, we assume that the supervisory review takes place at a specific "regulatory reporting dates", after banks issue deposits but before they draw their servicing costs (see Figure 5).[12] Given this timing, it is clear that the simple leverage requirement in Assumption 3 is not very effective. This is because the timing gives banks room to window–dress their balance sheet *ex ante*, by securitizing loans, before they are supervised, and to purchase loans back after they are supervised. We purposely introduce this "crack" in the regulation to study the effect of regulatory arbitrage on financial stability (see **?** and **?**).

The optimisation problem of the representative bank consists in maximizing its expected end–of–period t dividends with respect to d_t, s_t, $\mathbb{1}_t^i$, and ϕ_t (using (17)):

$$\begin{aligned}\max_{d_t, s_t} \mathbb{E}_{t-1} \Bigg(& \Psi_{t-1,t} \Big(r_t^s s_t - r_t^d d_t \\ & + \int_{\frac{\gamma}{r_t^\ell}}^1 \max_{\mathbb{1}_t^i, \phi_t} \Bigg((1 - \mathbb{1}_t^i) r_t^i (d_t + e_t - s_t) + \mathbb{1}_t^i \Big(q^\ell r_t^\ell (d_t + e_t - s_t + \phi_t) - r_t^i \phi_t \Big) \Bigg) \mathrm{d}\mu_\ell(q^\ell) \Big) \Bigg)\end{aligned} \tag{21}$$

subject to the incentive constraint (19), the regulatory constraint (20), and the securitization constraint (18). The first terms in (21) correspond to the profits relating to banks' first stage decisions —those taken before banks know their types. The first component is the return on securities, $r_t^s s_t$; the second is the payment of deposits, $r_t^d d_t$. The terms inside the integral correspond to the profits from the second stage decisions. The first component is the return, if bank q^ℓ sells its loan book to other banks, $r_t^i (d_t + e_t - s_t)$. The second component is the return if instead it purchases ϕ_t loans from other banks and then services $d_t + e_t - s_t + \phi_t$ loans. In this case, the bank

[12] This timing is meant to capture the regulator's limited information. In the context of our model, banking regulation is meant to put banks in the best possible position to access wholesale funding markets and cope with the *ex post* idiosyncratic q^ℓ–shocks.

must pay back the interbank loans, $r_t^i\phi_t$. The structure of the objective function in (21) reflects the bank's two–stage decision problem, whereby the bank first chooses (d_t, s_t) without knowing (z_t, γ) or its q^ℓ, and then chooses $(\mathbb{1}_t^i, \phi_t)$ knowing both (z_t, γ) and q^ℓ. Accordingly, we solve the maximization problem backward, starting with the choice of $\mathbb{1}_t^i$ and ϕ_t.

Second Stage: Choice of $\mathbb{1}_t^i$ and ϕ_t. From the linearity of (21), it is clear that bank q^ℓ purchases corporate loans on the secondary market whenever the net unit return on the leveraged funds is strictly positive, *i.e.* whenever $q^\ell r_t^\ell > r_t^i$. Hence,

$$\mathbb{1}_t^i = 1 \Leftrightarrow q^\ell \geq \overline{q}_t^\ell \equiv \frac{r_t^i}{r_t^\ell}. \tag{22}$$

The threshold $\overline{q}_t^\ell$ is the type of the worst banks that purchases loans. For a bank with $q^\ell > \overline{q}_t^\ell$, the first derivative of (21) with respect to ϕ_t is strictly positive, and the bank purchases as much loans as possible. As a result, (19) binds, which pins down the limit ϕ_t:

$$\phi_t = \frac{r_t^i - \gamma}{\gamma} x_t + \frac{r_t^s s_t + \kappa_t - r_t^d d_t}{\gamma}. \tag{23}$$

First Stage: Choice of d_t, and s_t. Given the second stage solutions (22) and (23), one can re–write (21) as:

$$\max_{d_t, s_t} \mathbb{E}_{t-1}\left(\Psi_{t-1,t}\left((1+\Delta_t)\left(r_t^i e_t + (r_t^s - r_t^i)s_t - (r_t^d - r_t^i)d_t\right) + \Delta_t \kappa_t\right)\right), \tag{24}$$

subject to the regulatory constraint (20) and the constraint (18), with

$$\Delta_t \equiv \frac{\left(1 - \mu_\ell\left(\overline{q}_t^\ell\right)\right)\left(\mathbb{Q}_t^\ell r_t^\ell - r_t^i\right)}{\gamma} \tag{25}$$

with

$$\mathbb{Q}_t^\ell \equiv \int_{\overline{q}_t^\ell}^1 q^\ell \frac{\mathrm{d}\mu_\ell(q^\ell))}{1 - \mu_\ell\left(\overline{q}_t^\ell\right)} > 1. \tag{26}$$

Let Λ_t denote the Lagrange multiplier associated with the regulatory constraint (20), and Γ the multiplier associated with the securitization constraint (18). Then, the first order conditions yield:

$$(1-\theta)\tau\Lambda_t = \mathbb{E}_{t-1}\left(\Psi_{t-1,t}(1+\Delta_t)r_t^i\right) - \mathbb{E}_{t-1}\left(\Psi_{t-1,t}(1+\Delta_t)\right) r_t^d \tag{27}$$

and

$$\Gamma = \tau\Lambda_t. \tag{28}$$

Relation (28) follows from shadow banks' no–arbitrage condition (16), and implies that banks securitize the maximum amount of loans possible ($\Gamma > 0$) whenever the regulatory constraint binds ($\Lambda_t > 0$). If $\Lambda_t > 0$, then $\Gamma > 0$ and $s_t = \theta(d_t + e_t)$. If $\Lambda_t = 0$, then $\Gamma = 0$, and two cases may arise.

If $e_t/(d_t+e_t) > \tau$, then banks comply with regulation without securitization, and are indifferent as to whether or not securitize their loans. By a continuity argument, we assume that if a bank is indifferent, then it does not securitize its loans and $s_t = 0$.[13] If $e_t/(d_t+e_t) < \tau$, then banks securitize loans up to what is necessary to comply with regulation, and $s_t = d_t + e_t - e_t/\tau$.

Since banks distribute their profits as dividends, the return on equity is given by:

$$r_t^e \equiv (1+\Delta_t)\left(r_t^i + (r_t^s - r_t^i)\frac{s_t}{e_t} - (r_t^d - r_t^i)\frac{d_t}{e_t}\right) + \Delta_t \frac{\kappa_t}{e_t}. \tag{29}$$

2.5 Decentralized General Equilibrium

A general equilibrium of this economy is defined as follows.

Definition 1 ***(Competitive General Equilibrium)*** *A competitive general equilibrium is:*

- *A sequence of prices,* $\mathcal{P}_t \equiv \{r_{t+j}^d, r_{t+j}^e, r_{t+j}^s, r_{t+j}^\ell, r_{t+j}^i\}_{j=0}^{j=\infty}$*;*
- *A sequence of quantities,* $\mathcal{Q}_t \equiv \{c_{t+j}, k_{t+j}, d_{t+j}, e_{t+j}, s_{t+j}, \ell_{t+j}, \phi_{t+j}\}_{j=0}^{j=\infty}$*,*

such that (i) for a given sequence of prices, $\mathcal{P}_t$*, the sequence of quantities,* $\mathcal{Q}_t$*, solves the optimization problems of the agents, and (ii) for a sequence of quantities,* $\mathcal{Q}_t$*, the sequence of prices,* $\mathcal{P}_t$*, clears the markets.*

Final Good Market. Denoting aggregate output by y_t, the final good market clears when

$$y_t = c_t + i_t,$$

where $i_t = k_{t+1} - (1-\delta)k_t$ and

$$\chi_t \equiv \underbrace{(1-\mathbb{Q}_t^\ell)r_t^\ell \ell_t + \left(\mathbb{Q}_t^\ell r_t^\ell - \gamma\right)\left(d_t + e_t - \left(1 - \mu_\ell\left(\overline{q}_t^\ell\right)\right)(\phi_t + x_t)\right)}_{\chi_t^\ell} + \underbrace{(1-\mathbb{Q}_t^d)r_t^d d_t}_{\chi_t^d} + \underbrace{(1-\mathbb{Q}_t^e)r_t^e e_t}_{\chi_t^e}$$

is the aggregate deadweight loss associated with transaction costs. The losses χ_t^d and χ_t^e are incurred by banks' depositors and shareholders, respectively, while χ_t^ℓ is incurred by the banks themselves. This latter cost is due to the servicing of loans by inefficient banks.

Interbank Loan Market. The interbank market clearing condition is given by:

$$\begin{cases} \left(1 - \mu_\ell\left(\overline{q}_t^\ell\right)\right)\phi_t = \mu_\ell\left(\overline{q}_t^\ell\right)x_t + s_t \text{ , if there is a unique solution with } \overline{q}_t^\ell > \dfrac{\gamma}{r_t^\ell} \\ \overline{q}_t^\ell = \dfrac{\gamma}{r_t^\ell}, \text{ otherwise.} \end{cases} \tag{30}$$

[13] This amounts to assuming that the bank incurs an infinitesimal securitization cost.

The above condition takes into account that banks are willing to lend only if $r_t^i \geq \gamma$. In the first line, the left side $\left(1-\mu_\ell\left(\overline{q}_t^\ell\right)\right)\phi_t$ is the aggregate demand for funding, and the right side, $\mu_\ell\left(\overline{q}_t^\ell\right)x_t + s_t$, is the aggregate supply. The latter is composed of the supply of funding from low q^ℓ banks as well as that from shadow banks, s_t. Using (17) and (16), one can re–write this condition as

$$\left(1-\mu_\ell\left(\overline{q}_t^\ell\right)\right)\left(\overline{q}_t^\ell + \frac{\kappa_t - r_t^d d_t}{r_t^\ell(d_t+e_t)}\right) = \frac{\gamma}{r_t^\ell}. \tag{31}$$

Given r_t^d and r_t^ℓ, it is easy to see that the expression on the left side is concave and hump–shaped in $\overline{q}_t^\ell$ over $\left[\frac{\gamma}{r_t^\ell}, 1\right]$, and admits one unique maximum at $\overline{q}_t^\ell = \overline{q}_t^{\max}$. Accordingly, a solution with $\overline{q}_t^\ell > \frac{\gamma}{r_t^\ell}$ exists and is unique if there is no solution at $\overline{q}_t^\ell = \frac{\gamma}{r_t^\ell}$, i.e. if

$$\left(1-\mu_\ell\left(\frac{\gamma}{r_t^\ell}\right)\right)\left(\frac{\gamma}{r_t^\ell} + \frac{\kappa_t - r_t^d d_t}{r_t^\ell(d_t+e_t)}\right) > \frac{\gamma}{r_t^\ell}. \tag{32}$$

Otherwise, the market clearing condition (30) admits multiple solutions, including one with $q_t^\ell = \frac{\gamma}{r_t^\ell}$. In that case, we assume that banks systematically fail to coordinate on the equilibria with the higher interbank rate and volume of transactions, and that the equilibrium solution is $q_t^\ell = \frac{\gamma}{r_t^\ell}$.[14] The first line in (30) characterizes the equilibrium with $r_t^i > \gamma$ and a high volume of transactions. We refer to this equilibrium as the "normal time" equilibrium. The second line characterizes an equilibrium with a low volume of transactions. In this equilibrium, the banks with $q_t^\ell > \frac{\gamma}{r_t^\ell}$ all borrow from the banks with $q_t^\ell = \frac{\gamma}{r_t^\ell}$. But their borrowing capacity is so limited that there is still excess supply of funding. The banks with $q_t^\ell = \frac{\gamma}{r_t^\ell}$ therefore must service the remaining loans on their balance sheet, even though this is socially inefficient. We call this equilibrium the "crisis time" equilibrium.

Role of Shadow Banking. Expression (30) points to the adverse effects of securitization in the model. Banks securitize their loans *ex ante*, before knowing their q^ℓ. Those include not only the low–q^ℓ, but also the high–q^ℓ banks, which will purchase loans back *ex post*. These two–way transactions by high–q^ℓ banks are inefficient. Shadow banks' supply of funds *ex post* exerts downward pressure on the equilibrium interbank rate, which crowds in low–q^ℓ banks, lowers the overall quality of the borrowers, and reduces banks' funding capacity (see (23)). As shadow banking impairs the reallocation of loans, it also raises the probability of a crisis.

3 Calibration

The calibration of the real sector of the economy is most standard. The capital elasticity in the production function, α, is set to 0.35, and the annual rate of depreciation of capital is set to $\delta = 0.06$.

[14] In an earlier work, we made the opposite assumption, ruling out coordination failures (see ?). Assuming coordination failures helps us calibrate the model, as in this case the model yields more crises (characterized by a low interbank rate).

The steady state level of TFP is normalized to one: $z = 1$. We set $\beta = 0.985$. The household's preferences are represented by the utility function

$$u(c_t) = \frac{c_t^{1-\frac{1}{\sigma}}}{1 - \frac{1}{\sigma}}.$$

We calibrate the financial parameters of the model for the US, based on the period 1988–2003. This period starts with the inception of the Basel Accord, and ends before its revamping in 2004 —which was also concomitant with the build up of financial imbalances that led to the great financial crisis of 2008. Over this period, the average leverage ratio (defined as equity over assets) for US banks was 8.1%, i.e. slightly above the Basel Accord minimum requirement of 8%. Accordingly, we will assume that the regulatory constraint was not binding, and set $\tau = 0.08$. As discussed earlier, this implies that there is no shadow banking activity in the steady state. For the sake of parsimony, we assume that the distributions of the transaction and loan servicing costs have the following form:

$$\mu_j(q) = q^{\lambda_j} \tag{33}$$

for $j \in \{d, e, \ell\}$. We model banks' strategic default cost, κ_t, as an iceberg cost proportional to the cash flow of the loan:

$$\kappa_t \equiv \epsilon r_t^\ell \hat{\ell}_t, \tag{34}$$

where $\hat{\ell}_t$ denotes the aggregate amount of loans in the economy (in the symmetric equilibrium, of course, $\hat{\ell}_t = \ell_t$). There are five financial parameters to calibrate: the average non–seizable cash flow, γ; the default cost parameter, ϵ; the parameters of the distributions of the financial transaction costs, λ_d and λ_e; and the parameter of the distribution of the servicing costs, λ_ℓ. Those parameters are jointly calibrated so that, in steady state, the model matches the targets listed in Table 1. The implied values of the parameters are reported in Table 2. The regulatory parameter, τ, is set to 8%. The limit on the amount of loans that a bank can securitize, θ, is arbitrarily set to 20%. Indeed, the regulatory constraint does not bind in the steady state, and therefore there is no securitization then. To get a sense of the sensitivity of the model to this parameter, we will experiment with several values.

4 Typical Path to Crises

We document the typical path to crises in the calibrated model.

4.1 Banking Glut versus Saving Glut

We distinguish three types of crises: "savings glut", "banking glut", and "both savings and banking glut". Savings glut crises are crises due to a savings glut. They are defined as crises that break

Table 1: Financial Targets

Model	Target	Data Sources
r^i	1.0194	FRED, Federal funds effective rate
r^ℓ	1.0428	FRED; Moody's seasoned Aaa corporate bond yield©
$e/(d+e-s)$	0.0810	US Financial Accounts; depository institutions (FL704194005.A-FL704190005.A)/FL704194005.A
$\sum_{i\in\{e,d\}}(1-\mathbb{Q}^i)r^i\frac{i}{a}$	0.0243	?
$(1-\mathbb{Q}^\ell)r^\ell\frac{\ell}{a}$	0.0100	FDIC Tables CB07 and CB09; banks' employee salaries and benefits to total assets
Proba of Crisis	0.0240	?

Table 2: Calibration

Parameter		Values
Parameterized		
Capital elasticity	α	0.300
Capital depreciation rate	δ	0.060
IES	σ	4.500
TFP (level)	z	1.000
Regulation	τ	0.080
Securitization limit	θ	0.100
Calibrated		
Discount factor	β	0.992
Non–seizable cash flow	γ	1.012
Distribution – $\mu_e(q^e)$	λ_e	0.501
Distribution – $\mu_d(q^d)$	λ_d	54.679
Distribution – $\mu_\ell(q^\ell)$	λ_ℓ	39.936
Default cost (level)	ϵ	1.549
Estimated		
TFP (persistence)	ρ^z	0.890
TFP (volatility)	σ^z	0.013
Default cost (persistence)	ρ^ϵ	–
Default cost (volatility)	σ^ϵ	–

out although there was no shadow banks in the five years that preceded it. We find that 56% of the crises in our economy are due to a pure savings glut. The remainder are due to either a pure banking glut, or to the concomitance of a banking and a savings glut. To tell the latter apart, we consider the sequence of shocks leading to those crises, and consider the state of the economy forty years before they begin. We simulate a counterfactual economy where shadow banking is prohibited, and simulate this economy using the same starting points and same sequences of shocks. We want to determine which crises are due to a banking glut. Those are the crises that would have broken out even in the absence of shadow banking. We find that 10% of the crises are due to both a savings and a banking glut. The remainder and the crises —34%— are thus due to shadow banking only.

4.2 Typical Path to Crises

To be completed…

Figure 2: Typical Path

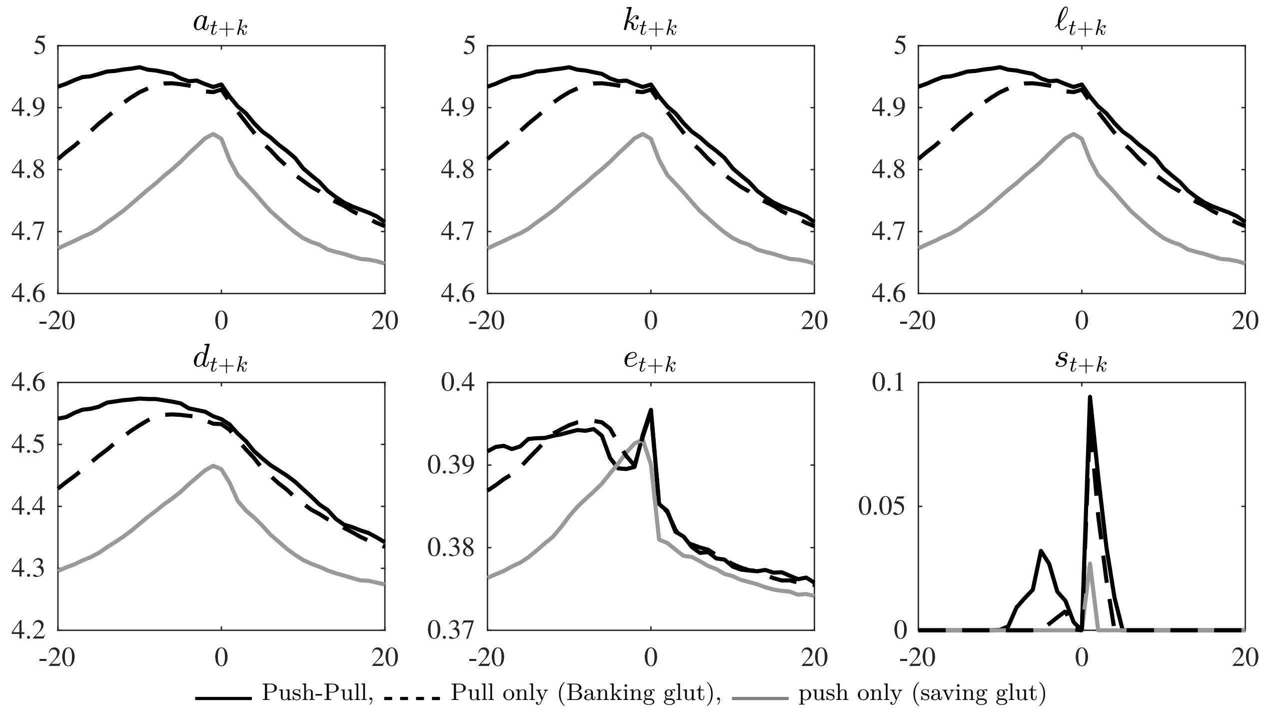


Figure 3: Typical Path

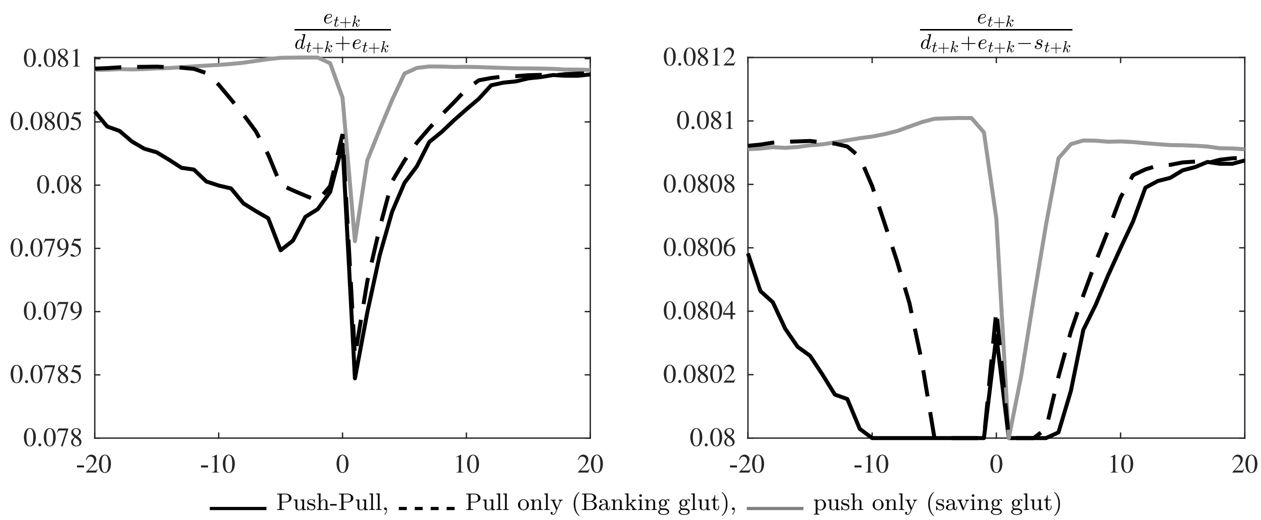
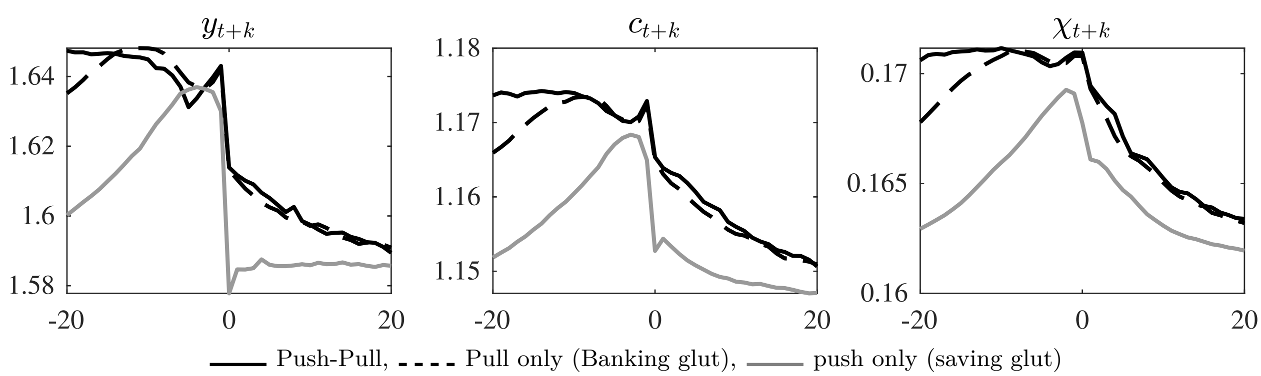


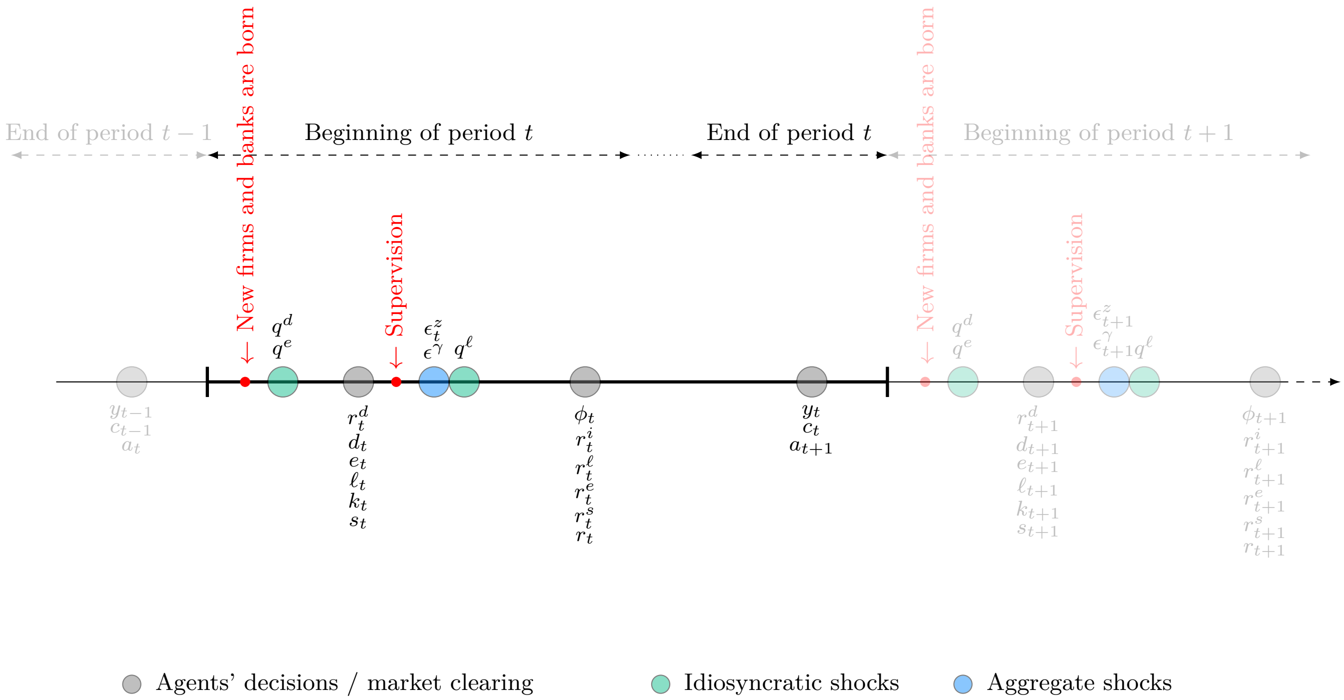
Figure 4: Typical Path



5 Appendix

5.1 Timeline

Figure 5: Shocks and Decisions in Period t



5.2 Equations of the Model

FIRMS

1. $y_t = z_t k_t^{\alpha} - \chi_t$
2. $k_t = \ell_t$
3. $r_t^{\ell} = \alpha z_t k_t^{\alpha-1} + 1 - \delta$

REPRESENTATIVE HOUSEHOLD

4. $\mathbb{E}_{t-1}\left(\Psi_{t-1,t} r_t\right) = 1$
5. $d_t = a_t \int_0^1 \mu_e \left(\min\left[1, q^d \frac{\mathbb{E}_{t-1}\left(\Psi_{t-1,t}\right) r_t^d}{\mathbb{E}_{t-1}\left(\Psi_{t-1,t} r_t^e\right)}\right]\right) \mathrm{d}\mu_d(q^d)$
6. $a_t = d_t + e_t$

BANKS

7. $d_t + e_t - s_t + \phi_t = \frac{r_t^i(d_t + e_t) + \kappa_t - r_t^d d_t}{\gamma}$
8. $(1-\theta)\tau\Lambda_t = \mathbb{E}_{t-1}\left(\Psi_{t-1,t}(1+\Delta_t) r_t^i\right) - \mathbb{E}_{t-1}\left(\Psi_{t-1,t}(1+\Delta_t)\right) r_t^d$
9. $\Gamma = \tau\Lambda_t$
10. $\Lambda_t\left(e_t - \tau(d_t + e_t - s_t)\right) = 0$
11. $\Gamma(\theta(d_t + e_t) - s_t) = 0$

    solution a: $\Lambda_t > 0$; $\Gamma > 0$; $s_t = \theta(d_t + e_t)$

    solution b: $\Lambda_t = 0$; $\Gamma = 0$; $s_t = d_t + e_t - \frac{e_t}{\tau}$ if $\frac{e_t}{d_t + e_t} < \tau$

    solution c: $\Lambda_t = 0$; $\Gamma = 0$; $s_t = 0$ if $\frac{e_t}{d_t + e_t} \geq \tau$
12. $\ell_t = a_t$
13. $\overline{q}_t^{\ell} \equiv \frac{r_t^i}{r_t^{\ell}}$

SHADOW BANK

14. $r_t^s = r_t^i$

MARKET CLEARING

15. $y_t = c_t + k_{t+1} - (1-\delta)k_t$

Normal Times (if $\left(1 - \mu_{\ell}\left(\frac{\gamma}{r_t^{\ell}}\right)\right)\left(\frac{\gamma}{r_t^{\ell}} + \frac{\kappa_t - r_t^d d_t}{r_t^{\ell}(d_t + e_t)}\right) \geq \frac{\gamma}{r_t^{\ell}}$)

16. $a_t = \left(1 - \mu_{\ell}\left(\overline{q}_t^{\ell}\right)\right)\left(\phi_t + d_t + e_t - s_t\right)$

Crisis Times (if $\left(1 - \mu_{\ell}\left(\frac{\gamma}{r_t^{\ell}}\right)\right)\left(\frac{\gamma}{r_t^{\ell}} + \frac{\kappa_t - r_t^d d_t}{r_t^{\ell}(d_t + e_t)}\right) < \frac{\gamma}{r_t^{\ell}}$)

16. $\overline{q}_t^{\ell} = \frac{\gamma}{r_t^{\ell}}$

Definitions

17. $\mathbb{Q}_t^d \equiv \frac{a_t}{d_t} \int_0^1 q^d \mu_e \left(\min \left[1, q^d \frac{\mathbb{E}_{t-1}\left(\Psi_{t-1,t}\right) r_t^d}{\mathbb{E}_{t-1}\left(\Psi_{t-1,t} r_t^e\right)} \right] \right) \mathrm{d}\mu_d(q^d)$
18. $\mathbb{Q}_t^e \equiv \frac{a_t}{e_t} \int_0^1 q^e \mu_d \left(\min \left[1, q^e \frac{\mathbb{E}_{t-1}\left(\Psi_{t-1,t} r_t^e\right)}{\mathbb{E}_{t-1}\left(\Psi_{t-1,t}\right) r_t^d} \right] \right) \mathrm{d}\mu_e(q^e)$
19. $\mathbb{Q}_t^\ell \equiv \int_{\overline{q}_t^\ell}^1 q^\ell \frac{\mathrm{d}\mu_\ell(q^\ell))}{1 - \mu_\ell\left(\overline{q}_t^\ell\right)}$
20. $\Psi_{t,t+1} \equiv \beta \frac{u'(c_{t+1})}{u'(c_t)}$
21. $\chi_t \equiv (1 - \mathbb{Q}_t^e) r_t^e e_t + (1 - \mathbb{Q}_t^d) r_t^d d_t + (1 - \mathbb{Q}_t^\ell) r_t^\ell a_t + \left(\mathbb{Q}_t^\ell r_t^\ell - \gamma\right) \left(a_t - \left(1 - \mu_\ell\left(\overline{q}_t^\ell\right)\right)(\phi_t + d_t + e_t - s_t)\right)$
22. $\Delta_t \equiv \frac{\left(1 - \mu_\ell\left(\overline{q}_t^\ell\right)\right)\left(\mathbb{Q}_t^\ell r_t^\ell - r_t^i\right)}{\gamma}$
23. $r_t a_t \equiv \mathbb{Q}_t^d r_t^d d_t + \mathbb{Q}_t^e r_t^e e_t$
24. $r_t^e \equiv (1 + \Delta_t) \left(r_t^i - (r_t^d - r_t^i) \frac{d_t}{e_t} \right) + \Delta_t \frac{\kappa_t}{e_t}$
25. $\overline{q}_t^{\max} \equiv \arg\max_{\overline{q}_t^\ell} \left(1 - \mu_\ell\left(\overline{q}_t^\ell\right)\right) \left(\overline{q}_t^\ell + \frac{\kappa_t - r_t^d d_t}{r_t^\ell (d_t + e_t)} \right)$
26. $\kappa_t \equiv \epsilon r_t^\ell \ell_t$

There are 26 equations and 26 unknowns:

y	c	k	a	e	d	ϕ	ℓ	s	Λ	Γ	χ
1	2	3	4	5	6	7	8	9	10	11	12

$\mathbb{Q}^\ell$	$\mathbb{Q}^d$	$\mathbb{Q}^e$	r	r^i	r^s	r^e	r^d	r^ℓ	Ψ	Δ	$\overline{q}^\ell$	$\overline{q}^{\max}$	κ_t
13	14	15	16	17	18	19	20	21	22	23	24	25	26

There are 9 parameters to calibrate: 4 from the real side, and 5 from the financial side of the economy:

z	α	β	δ	λ_d	λ_e	λ_ℓ	γ	ϵ
1	2	3	4	5	6	7	8	9

Parameters τ and θ are parameterized.

5.3 Banks' Balance Sheets

Figure 6: Banks' Balance Sheets at the Regulatory Reporting Dates

Shadow banks	
s_t	s_t

Firms	
k_t	ℓ_t

Banks	
x_t	d_t
s_t	e_t

Households	
d_t	a_t
e_t	