EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1729 / SEPTEMBER 2014

FLIGHT TO LIQUIDITY AND THE GREAT RECESSION

Sören Radde

In 2014 all ECB publications feature a motif taken from the €20 banknote.

Acknowledgements
I gratefully acknowledge helpful comments by Markus Brunnermeier, Wouter Den Haan, Monique Ebell, Christoph Große Steffen, Hans-Peter Grüner, Frank Heinemann, Matej Marinc, Simon Junker and an anonymous referee as well as participants at the DIW Seminar on Macro and International Economics, the Annual Conference of the DFG Collaborative Research Center 649 "Economic Risk", the 9th Workshop on Money, Banking, and Financial Markets at Heinrich-Heine-University Düsseldorf, the research seminar at De Nederlandsche Bank, the 2012 European Workshop in Macroeconomics at Oesterreichische Nationalbank, the 4th International IFABS Conference at the University of Valencia, the 2012 EEA Congress at the University of Málaga and the Annual Conference of the Verein für Socialpolitik at the University of Göttingen. I thank the DIW Berlin Graduate Center for providing financial support for this project.

Sören Radde
European Central Bank, German Institute for Economic Research
and Technische Universität Berlin; e-mail: soeren.radde@ecb.europa.eu

Address	Kaiserstrasse 29, 60311 Frankfurt am Main, Germany
Postal address	Postfach 16 03 19, 60066 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Internet	http://www.ecb.europa.eu

ISSN	1725-2806 (online)
ISBN	978-92-899-1137-5
EU Catalogue No	QB-AR-14-103-EN-N (online)

Abstract

This paper argues that counter-cyclical liquidity hoarding by financial intermediaries may strongly amplify business cycles. It develops a dynamic stochastic general equilibrium model in which banks operate subject to agency problems and funding liquidity risk in their intermediation activity. Importantly, the amount of liquidity reserves held in the financial sector is determined endogenously: Balance sheet constraints force banks to trade off insurance against funding outflows with loan scale. A financial crisis, simulated as an abrupt decline in the collateral value of bank assets, triggers a flight to liquidity, which strongly amplifies the initial shock and induces credit crunch dynamics sharing key features with the Great Recession. The paper thus develops a new balance sheet channel of shock transmission that works through the composition of banks' asset portfolios.

Keywords: macro-finance; funding liquidity risk; liquidity hoarding; bank capital channel; credit crunch.

JEL classification: E22; E32; E44.

Non-technical summary

The recent financial crises and ensuing recessions in the United States and the euro area have highlighted the central role of financial intermediaries in the transmission and amplification of financial sector shocks to the real economy. A pronounced maturity mismatch between funding and assets exposed banks to roll-over risk in their refinancing operations. This risk materialised as collateral pools used on markets for secured short-term refinancing dried up and counter-party risk was perceived to rise. Banks responded by fleeing into liquid assets, eventually crowding out lending to the non-financial sector, thus transmitting financial sector stress to the real economy.

While the finance literature has long recognised funding liquidity risk and precautionary liquidity hoarding as important financial sector vulnerabilities, the macroeconomic effects of a broad-based flight to liquidity in the banking sector have not been explicitly explored. In this paper, I study the trade-off that determines banks' portfolio choice between liquid reserves and loans to the private sector and its impact on macroeconomic stability. Banks react to financial-sector specific shocks by shoring up their liquidity buffers, which acts as a powerful amplification mechanism leading to a sharp economic downturn.

The framework used for the analysis is an extension of the canonical real business cycle model with financial intermediaries and funding liquidity risk at the bank level. Banks finance loans to capital producers with deposits received from households. Households may withdraw a fraction of their deposits during the life-time of loan projects, exposing banks to roll-over risk. To insure against such risk, intermediaries hold liquidity reserves in the form of contingent credit lines with a mutual fund. However, liquidity risk cannot be fully diversified due to a moral hazard problem between banks and the mutual fund. Therefore, funding outflows in excess of liquidity buffers result in the termination of loan projects, with the mutual fund collecting the liquidation proceeds as its collateral. Given limited inside and outside funding, banks thus need to trade off insurance against funding liquidity risk via liquidity buffers with profitable lending to capital producers as part of their portfolio choice.

The model demonstrates how a financial-sector specific shock that reduces the collateral value of bank loans induces a flight to liquidity and lending squeeze. Lower collateral values make the liquidation of loan projects more costly for the outside investor, such that the threat to not roll over banks' funding becomes less effective. Banks can thus extract more profit from a given loan project and prefer its continuation. Therefore, they hoard liquidity to withstand larger funding outflows to the detriment of their lending scale. The resulting credit crunch strongly amplifies the initial losses incurred in the financial sector due to the decline in bank loans' collateral value. The model is able to explain a number of salient features observed during the recent financial crisis, such as the counter-cyclical flight to liquidity as well as pro-cyclical lending and leverage.

Finally, the paper shows that price rigidities may interact with financial frictions to strongly exacerbate the amplification of financial sector shocks. With sluggish price adjustment, the recession triggered by a collateral shock is associated with falling inflation due to shrinking profit

margins, which result in wage cuts and a decline in private consumption. These effects are stronger the less aggressively the monetary authority is able to cut its policy rate in order to offset the impact of inflation on real variables. This feature is relevant for central banks as it hints at a potentially important source of distortions introduced by the zero lower bound.

1 Introduction

The recent financial crisis and the ensuing Great Recession started with an abrupt surge in uncertainty about the value of assets on financial intermediaries' balance sheets. As investors became worried about counter-party risk and the quality of collateral pools, financial institutions found it difficult to roll over short-term debt. In fact, several markets for short-term refinancing experienced runs: Asset-backed commercial paper became illiquid in late 2007, followed by a freeze in the unsecured interbank-market after the demise of Lehman Brothers (Brunnermeier, 2009; Heider, Hoerova, and Holthausen, 2009). Similarly, average haircuts on collateral assets in repurchase agreements (repo) rose from zero to 45% within the span of one year, effectively withdrawing \$1.2 trillion in funding from the repo market (Gorton and Metrick, 2010, 2012; Duffie, 2010).

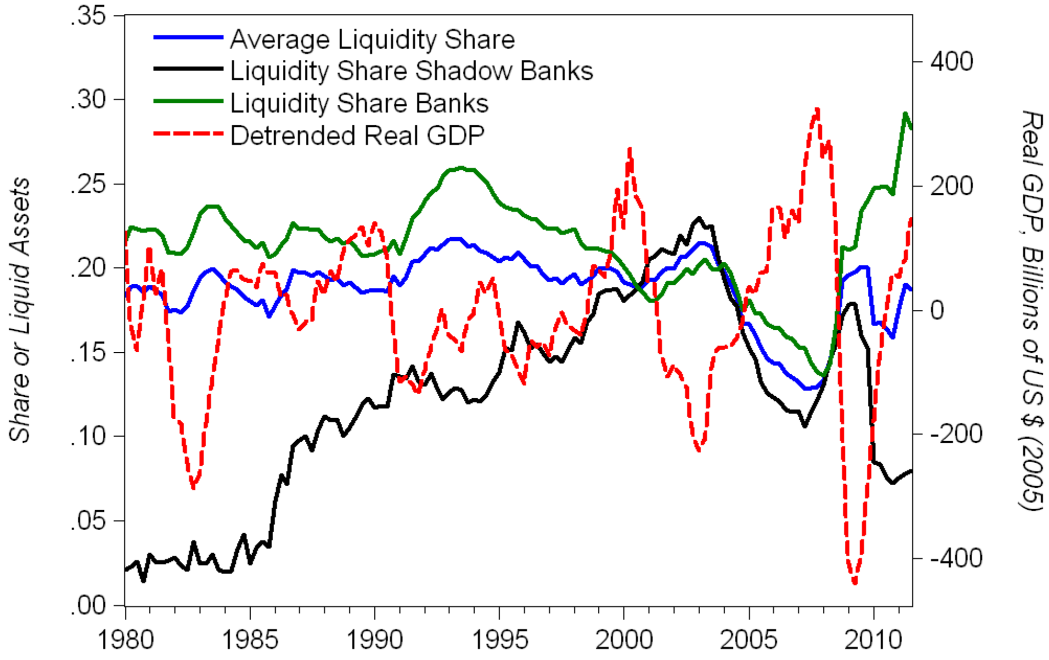
In response to the downward spiral of plummeting collateral values and rising funding liquidity risk, financial institutions took to hoarding liquid assets in order to reduce the maturity mismatch on their balance sheets. In the US, the flight to liquidity episode started in late 2007, manifested in a starkly rising share of liquid assets in total balance sheet size (see Figure 1). In fact, liquidity shares had been counter-cyclical since the early 2000s both in the traditional and in the shadow banking sector, with a contemporaneous cross-correlation of -0.46 and -0.40, respectively.[1] However, the hoarding of liquid reserves locked up funds otherwise available for investment into riskier assets. This curtailed the lending capacity of the financial sector and eventually impaired non-financial firms' access to external financing, thus propagating financial sector stress to the real economy.

In order to capture these events, this paper develops a model that features flight to liquidity in the banking sector as the key amplification and propagation mechanism of financial sector stress. Financial intermediaries operate subject to moral hazard problems in their monitoring activity and funding liquidity risk. They trade off the amount of liquidity reserves to hold as insurance against funding outflows with the amount of funds available for lending. Aggregate shocks to the collateral value of bank assets trigger a flight to liquidity in the sense of higher insurance against short-term funding risk. This amplifies the initial shock and induces credit crunch dynamics sharing key features with the Great Recession.

The model extends the canonical real business cycle model with financial intermediation and liquidity risk at the bank level. In particular, banks intermediate funds between savers and capital producers and provide monitoring services for which they bear a cost. This private cost creates an agency problem between outside investors and banks: Banks need to retain an equity stake in their loans in order to receive a sufficient fraction of the return on lending that compensates them

[1] I define liquid assets as the sum of checkable deposits and currency, cash and reserves at the Federal Reserve, Treasury securities as well as agency- and securities backed by Government-Sponsored-Enterprise (GSE). Of course, if these are truly liquid assets, they are expected to retain their value during a downturn, while prices of riskier assets would fall. Thus, the value of liquid assets relative to total balance sheet size would mechanically increase. However, the fact that liquidity buffers were not adjusted downwards suggests that a flight to liquidity occurred and banks' willingness to lend declined.

Figure 1: Share of Liquid Assets of Banks and Market-based Intermediaries



Notes: The liquidity share is computed as the sum of checkable deposits and currency, cash and reserves at the Federal Reserve, Treasury securities, agency and GSE-backed securities relative to total assets of the respective institutions. Source: US Flow of Funds (Federal Reserve)

for their monitoring services. As banks retain a fraction of the return on loans, the agency problem drives a wedge between the total return on loans and the amount that is pledgeable to outside investors. Building on Holmström and Tirole (1998), liquidity shocks arrive during the life-time of loan projects that require the input of additional resources. Here, such shocks are modelled as withdrawals of external funds, which are idiosyncratic at the bank level. Economically, they amount to rollover risk arising from a maturity mismatch between bank-assets and bank-funding. Limited pledgeability of loan returns constrains the funds that can be attracted to refinance these outflows, such that even projects with a higher continuation than liquidation value may have to be terminated.

Anticipating their financial constraints at the lending and at the refinancing stage, banks need to decide on how to optimally allocate their inside and outside funding between loans and liquidity reserves simultaneously. Given limited financial resources, earmarking funds as liquidity reserves decreases the scale of loans that banks can extend before liquidity shocks arrive. This trade-off implies that banks choose not to fully insure against liquidity risks. As a consequence, funding outflows cannot be entirely diversified despite being idiosyncratic. In particular, funding withdrawals in excess of liquidity reserves lead to the termination and inefficient liquidation of investment projects by the outside financiers.

Following evidence on rising haircuts in repo transactions for secured short-term finance, I introduce a shock to the liquidation - or collateral - value of bank assets as a novel type of aggregate risk.[2] With lower collateral values, the liquidation of loan projects becomes more costly for outside investors. Banks can thus extract more resources after liquidity shocks have occurred, which makes liquidity reserves more compelling relative to the initial scale of loans. The flight to liquidity unleashes a powerful amplification mechanism as higher liquidity reserves crowd out funds for initial bank lending (bank lending channel). These dynamics stand in sharp contrast to a frictionless economy where such crowding-out would not occur. Extending the model with nominal frictions, I also demonstrate how these interact with financial frictions to exacerbate the amplification from and recessionary impact of a liquidity crisis.

The contribution of the paper is twofold: First, it introduces funding liquidity risk arising from a maturity mismatch between financial assets and external finance into a dynamic stochastic general equilibrium framework. Second, it explains how shocks operating directly on the balance sheets of financial intermediaries are amplified due to an endogenous increase in aggregate liquidity demand which emerges from the interaction between agency costs and funding liquidity risk. This adds to the literature on the balance sheet channel of shock transmission. However, amplification works through the endogenous composition of balance sheets, i.e. the choice between insurance against liquidity risk and lending scale, rather than fluctuations in the net worth of borrowers as in the financial accelerator literature. The model can explain a number of salient features observed during the recent financial crisis, such as the counter-cyclical flight to liquidity phenomenon as well as pro-cyclical lending and leverage.

1.1 Related Literature

This paper contributes to the growing body of literature on macro-financial linkages. It builds on two distinct strands of research. The first analyses financial frictions as the source of business cycle fluctuations. At the heart of this research is the balance sheet channel as surveyed by Bernanke and Gertler (1995), i.e. the amplification and propagation of business cycles due to a financial accelerator mechanism arising from the feedback between borrowing constraints and asset fire sales. Theoretical research in this area focuses on agency frictions between borrowers and lenders. Carlstrom and Fuerst (1997), Bernanke and Gertler (1989), Bernanke, Gertler, and Gilchrist (1999) and more recently Christiano, Motto, and Rostagno (2014) have embedded the costly-state-verification framework developed by Townsend (1979) in relation to financial contracts into business cycle models to study the dynamic impact of such agency costs. Other studies, such as Kiyotaki and Moore (1997), Gertler and Karadi (2011) and Gertler and Kiyotaki (2011), rely on limited or costly enforcement of financial contracts. Holmström and Tirole (1997) study an incentive model of financial intermediation where both firms and banks are capital constrained. The business cycle implications of this bank capital channel are analysed by Meh and Moran

[2] The liquidation value of a loan measures its resale value. This corresponds to the concept of *market liquidity* as defined in Brunnermeier, Eisenbach, and Sannikov (2012).

(2010), which is closely related to this paper.

However, the literature discussed so far does not accommodate the notion of endogenous liquidity demand. I introduce this feature following a second strand of literature initiated by Holmström and Tirole (1998). These authors develop a finite-horizon framework which motivates demand for corporate liquidity reserves with uncertain reinvestment needs during the lifetime of investment projects. Kato (2006) incorporates this structure into a dynamic general equilibrium setting to analyse the business cycle dynamics that result from liquidity risk at the corporate level. His model is able to replicate firms' counter-cyclical dependence on external finance and the hump-shaped response of output to shocks observed in the US. Covas and Fujita (2010) expand this analysis by adding regulatory capital requirements in the banking sector.

In Kurlat (2013) and Bigio (2013), liquid assets are needed to relax financial constraints in financing capital investment and working capital loans. Brunnermeier and Sannikov (2014) merge the financial accelerator framework of Kiyotaki and Moore (1997) and Bernanke, Gertler, and Gilchrist (1999) with the financial intermediation model of Holmström and Tirole (1997) to study asset fire sale episodes that endogenously decrease the market value of bank assets. These frameworks have in common that they focus on asset market liquidity, while funding liquidity risk is not considered. Similarly, they do not account for the flight to liquidity observed in the financial sector during the crisis.

The present paper merges the literature on the role of bank capital in the business cycle with the role of liquidity for relaxing financial constraints. However, I depart from previous research in a number of ways. First, liquidity risk is introduced at the bank level. Second, liquidity risk is modelled as funding risk. Third, the collateral value of liquidated investment projects for banks' financiers is assumed to be non-zero.

Shocks related to financial markets are seen as potential drivers of the 2007-09 financial crisis. Gertler and Karadi (2011) model a shock to capital quality, which depresses the value of bank assets and triggers fire sales due to a leverage constraint imposed on banks. The resulting credit crunch drives the economic downturn. Meh and Moran (2010) investigate the business cycle consequences of a direct shock to bank capital, which in their model has limited recessionary impact. Other studies focus on shocks affecting the collateral value of financial assets. Kurlat (2013) and Bigio (2013) study the impact of information shocks that decrease the market liquidity of assets due to asymmetric information about asset quality, thus tightening the financing frictions in their respective models. In a model of endogenous information production on financial asset quality, Gorton and Ordoñez (2014) impose a shock on the perceived value of collateral assets whose true value is opaque. This leads to a break-down of short-term debt markets relying on high-quality collateral. All of these scenarios have in common that the collateral value of assets held by financially constrained agents drops due to some exogenous disturbance.

Following these contributions, I introduce a shock to the liquidation value of bank assets as a new source of aggregate risk. This shock is intended to mimic the sudden drop in collateral values at the onset of the financial crisis described by Gorton and Metrick (2010, 2012). It is a parsimonious, reduced-form approach to capturing asset devaluations due to information, agency or search frictions. Moreover, it is more subtle than some of the financial-sector specific shocks introduced in the recent literature. For instance, unlike Meh and Moran (2010) the shock does not operate directly on banks' equity capital, rather it exploits their balance sheet constraints; unlike Gertler and Karadi (2011) it only assumes that the value of liquidated projects, but not that of successful projects decreases.

The remainder of this paper is organized as follows. Section 2 develops the model. The baseline calibration of the model and the aggregate shock processes as well as simulation results are discussed in section 3. This section also offers insights into the role of nominal in addition to financial frictions. Section 4 concludes.

2 The Model

2.1 The environment

Consider an economy populated by five types of agents: a continuum of agents with unit mass comprising a large fraction η^h of consumers (households) and a small fraction $\eta^b = 1 - \eta^h$ of bankers; a mutual fund; and capital good and final good producers. Time is discrete and infinite ($t = 0, 1, 2, ...$).

There are two goods in the economy. Final or consumption goods are produced in a frictionless, competitive market. Capital goods are produced by entrepreneurs who convert final goods into capital goods subject to idiosyncratic risk of failure. Entrepreneurs are financed and monitored by banks, who suffer from agency problems towards their outside financiers and funding liquidity risk due to a maturity mismatch between their assets and liabilities. Banks receive funding in the form of deposits from a mutual fund, which intermediates the savings of the household sector. This fund also extends contingent credit lines that banks can draw on to partially buffer funding outflows. Each of these agents is now described in detail.

2.2 Capital Good Producers

Capital good producers manage investment projects in order to produce new capital goods. They have access to a stochastic constant returns to scale technology converting i_t units of consumption goods into Ri_t units of capital, if successful, and zero otherwise.[3] By exerting effort, entrepreneurs can ensure a high probability of success, π_H. If they shirk the success probability of investment projects is reduced to $\pi_L < \pi_H$, but entrepreneurs earn a private benefit. This creates a moral

[3]The capital production technology is assumed to exhibit constant returns to scale in order to stay as close as possible to a frictionless real business cycle model. This avoids introducing distortions that would obscure the main mechanism analysed in this paper.

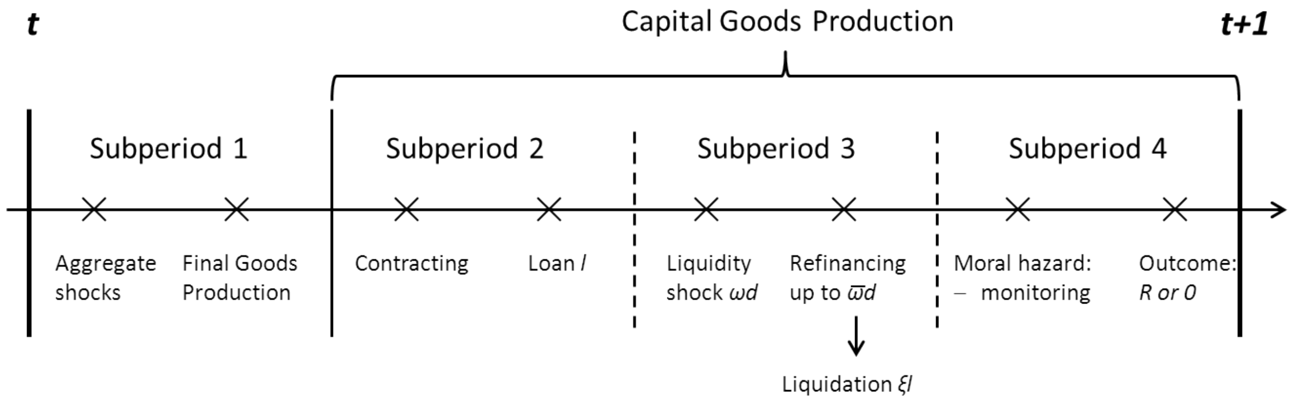


Figure 2: Timing

hazard problem: In order to exert effort, capital producers have to be either compensated for foregoing private benefits or monitored.

Any compensation for unmonitored entrepreneurs would reduce the return that an outside investor, who strives to induce effort, could extract from entrepreneurs. However, capital producers enter a close relationship with their lending banks when seeking external finance for their investment projects following Holmström and Tirole (1997). In this relationship, banks are assumed to have the capacity to detect shirking via some monitoring technology. Monitored entrepreneurs are prevented from shirking without having to be compensated. Hence, monitoring by banks eliminates the moral hazard problem in capital producers and ensures that entrepreneurs can channel all returns from investment projects to their lending banks.

2.3 Financial Sector

Timing. The intra-period timing assumptions are crucial for the set-up of the financial sector. Every period is divided into four subperiods (Figure 2). In the first subperiod, aggregate shocks occur and production of final goods takes place. Capital goods production extends over the last three subperiods: In subperiod 2, financial contracts are negotiated between banks and their outside investor (the mutual fund) and loans are extended to capital producers. Each bank extends a unique loan to an entrepreneur that is used to finance investment on a one-to-one basis, i.e. $i_t = l_t$.[4] After loans have been committed, liquidity shocks occur in subperiod 3 and loans need to be refinanced or liquidated. In the final subperiod, surviving loan projects have to be monitored by banks and new capital goods are produced by successful investment projects. Finally, all parties are paid according to their contracts.

Liquidity shocks. At the initial loan-financing stage in subperiod 2, banks receive outside funding in the form of intra-period deposits d_t, which are collected from households and channelled

[4] Due to the lack of entrepreneurial capital, bank loans effectively amount to outside equity stakes in entrepreneurial projects.

to the banking sector by the mutual fund. Liquidity shocks take the form of withdrawals of deposits that are idiosyncratic at the bank level and uncertain at the time of contracting. The funding outflows are proportional to the amount of external finance held by banks, i.e. a fraction ω of deposits d_t is withdrawn, where the random variable $\omega \in [0, U_H]$ is distributed according to the cumulative distribution function $F(\omega)$ and density $f(\omega)$. Withdrawals flow back to the household sector and are used for consumption purposes.[5] Liquidity shocks may be thought of as rollover risk arising from a maturity mismatch between bank-assets and bank-funding. Assuming such idiosyncratic shocks at the bank-level serves as a short-cut for modelling heterogeneity in banks' funding structures.

First-best refinancing threshold. Suppose that monitoring was costless for banks, such that the entire return on their loans could be pledged to outside investors. Initial loans could then be financed entirely with outside funding from the mutual fund in subperiod 2. In subperiod 3, banks need to refinance funding outflows ωd_t in order to continue their loan project. In the absence of monitoring costs, the mutual fund would be willing to provide this additional funding up to the amount that makes it indifferent between the continuation of a loan project and its liquidation. The first-best refinancing or liquidity threshold after deposit withdrawals have materialised thus satisfies

$$\underbrace{q_t \pi_H R l_t - \bar{\omega}_1 d_t}_{\text{continuation}} = \underbrace{q_t \xi_t^* l_t}_{\text{liquidation}}$$

where $\pi_H R l_t$ is the (expected) return to a successful, monitored loan project in terms of capital goods, q_t is the price of capital in terms of consumption goods and ξ_t^* is the exogenously determined fraction of the initial loan scale that the outside investor can salvage in case of liquidation. For any $\bar{\omega}_t > \bar{\omega}_1$, loan projects have a lower continuation than liquidation value even in the absence of monitoring costs and the mutual fund would prefer liquidation. With monitoring costs, the first-best refinancing threshold cannot be implemented and liquidation will ensue following much smaller occurrences of the liquidity shock.[6]

Moral hazard. By enforcing entrepreneurial effort through their monitoring activity, banks eliminate the agency friction in the capital production process. Monitoring services are assumed to involve private costs, such that the relationship between banks and their financier is affected by a moral hazard problem. In particular, monitoring banks incur costs in terms of final goods which are proportional to project size, i.e. μl_t. They must earn a minimum return in order to cover this cost and thus cannot pledge the entire return on their loan to the mutual fund. Therefore, neither initial loans nor liquidity needs at the refinancing stage can be fully financed with outside

[5] An aggregate outflow of resources from the financial sector can be motivated by resource losses in financial intermediation arising from costs faced by outside investors. Institutional investors may, for instance, have to put up additional resources to investigate banks' soundness or monitor their behaviour while funding them. Such costs would need to be covered by a premium on external funding arising while the relationship is ongoing.

[6] For further details see Proposition 1.

funding. As a result, banks need to provide some inside funding in the form of bank capital a_t, which is accumulated through retained earnings.[7]

Key mechanism. Anticipating their financial constraints both at the initial contracting and at the refinancing stage, banks need to decide on how to optimally allocate their available inside and outside funds between initial loans and liquidity reserves (plus expected monitoring costs) simultaneously. This portfolio allocation also needs to ensure a non-negative expected return for the fund (investor participation constraint) and induce monitoring efforts by banks (incentive compatibility).[8] The financial contract that achieves these objectives is discussed in section 2.3.1. Given limited financial resources, earmarking funds as liquidity reserves reduces the scale of loans that banks can initially extend. This trade-off is the key mechanism of the model.

Liquidity reserves. The optimal amount of liquidity reserves that follows from the trade-off with banks' loan scale is the central element spelled out by the financial contract between banks and their outside investor. Liquidity reserves take the form of uncommitted funds that banks leave with the mutual fund in subperiod 2, which entitle them to draw a credit line after having been hit by a funding outflow in subperiod 3. Banks choose the optimal liquidity buffer by capping their contingent credit line at some upper threshold $\bar{\omega}_t < \bar{\omega}_1$. This will allow them to obtain up to $\bar{\omega}_t d_t$ in additional funding from the mutual fund in case of liquidity shocks. Since these shocks are uncertain at the time of contracting, the corresponding reserve that is set aside in subperiod 2 amounts to the expected liquidity outflow given the optimal threshold, i.e. $\left(\int_0^{\bar{\omega}_t} \omega \, \mathrm{d}F(\omega)\right) d_t$.[9]

Refinancing in excess of the liquidity threshold would violate banks' incentive to engage in costly monitoring. Hence, projects with funding outflows $\omega > \bar{\omega}_t$ are liquidated by the outside investor. Given the distribution of the liquidity shock, the *ex ante* probability of an investment project to survive the funding shock is $\int_0^{\bar{\omega}} f(\omega) \, \mathrm{d}\omega = F(\bar{\omega}) < F(\bar{\omega}_1)$.[10] This shows that liquidity shocks - although idiosyncratic at the bank-level - cannot be fully diversified due to the agency problem between banks and the mutual fund. Accordingly, some investment projects with positive net present value, i.e. continuation values in excess of liquidation values, have to be terminated even when banks accumulate liquidity reserves.

No partial liquidation. Loan projects are assumed to be indivisible given banks' role as relationship lenders. Banks acquire specific information about their debtors, which is costly, because it takes time to build. Partial liquidation of loan projects would disrupt this relationship and hence worsen both the monitoring incentives and ability of banks. Buyers of tranches of other intermediaries' loan portfolios could at best substitute imperfectly for the monitoring services provided by the original lender. Thus, partial liquidation would aggravate the moral hazard problem

[7] For details on the accumulation of bank capital see section 2.3.3.

[8] I rule out equilibria in which effort is not induced for bankers.

[9] This arrangement between a bank and the mutual fund is akin to an insurance contract with the liquidity buffer resembling the insurance premium. The financing of the actual credit lines that are granted by the mutual fund after liquidity shocks materialise is discussed in section 2.3.2.

[10] Note that liquidation only occurs once liquidity buffers have been exhausted entirely. The liquidation value ξ_t^* is hence independent of the funding characteristics of banks and the mutual fund has no resources other than unfinished loan projects to draw on.

between banks and depositors, increase further the wedge between total and pledgeable returns further and thus make loan financing more difficult. For these underlying structural reasons, partial liquidation is ruled out.[11]

2.3.1 Intra-period Financial Contract

Constrained problem. The optimal financial contract is a set $\{l_t, d_t, R_t^b, R_t^h, \bar{\omega}_t\}$ which specifies the level of loans l_t, the amount of deposits d_t, the distribution of per unit project return R to banks, R_t^b, and households, R_t^h, as well as the threshold level of the liquidity shock, $\bar{\omega}_t$, which banks can accommodate by tapping into their liquidity buffer. General equilibrium effects have an impact on the financial contract through the beginning-of-period relative price of capital q_t, the previously accumulated capital of banks a_t and the stochastic liquidation value ξ^*. At the time of contracting, these are, however, exogenous.

Since the contracting problem takes place within a period, time subscripts are omitted in the description of the optimal contract. Formally, the contract maximises banks' expected return from loans to entrepreneurs subject to incentive compatibility, participation, and feasibility constraints:

$$\max_{\{l,d,R^b,R^h,\bar{\omega}\}} qF(\bar{\omega})\pi_H R^b l$$

$$\text{s.t.}$$

$$qF(\bar{\omega})\pi_H R^b l - F(\bar{\omega})\mu l \geq qF(\bar{\omega})\pi_L R^b l \tag{1}$$

$$q\left[F(\bar{\omega})\pi_H R^h + (1-F(\bar{\omega}))\xi^*\right] l \geq d \tag{2}$$

$$d + a \geq (1 + F(\bar{\omega})\mu)\, l + \left(\int_0^{\bar{\omega}} \omega \,\mathrm{d}F(\omega)\right) d \tag{3}$$

$$R = R^b + R^h \tag{4}$$

The objective function accounts for the fact that the probability of successfully executing a project of scale l is $F(\bar{\omega})\pi_H$, since the *ex ante* probability of a non-excessive liquidity shock is $F(\bar{\omega})$, and the probability of yielding non-zero output is π_H. As indicated by their incentive compatibility constraint (1), bankers need to be compensated with $R^b \geq \frac{\mu}{q(\pi_H - \pi_L)}$ in order to monitor entrepreneurs. The share of loan returns that banks need to retain captures the severity of the moral hazard problem with respect to banks' outside financiers and drives the crucial wedge between the full and the pledgeable return to loans as $R - R^h = R^b$.[12] Equation (2) is the participation constraint of the

[11] Fixed monitoring costs would also be an obstacle to partial liquidation: Every time a loan tranche was sold, buyers would have to incur these fixed costs, which would amount to significant efficiency losses in financial intermediation and render partial liquidation unattractive. As discussed by Holmström and Tirole (1998), partial liquidation is also not an optimal policy in this setup. Agents prefer full continuation as long as they can accommodate liquidity shocks and full liquidation if shocks exceed their liquidity buffers.

[12] The incentive compatibility constraint of banks tightens when monitoring costs rise or when the probability differential between monitored and non-monitored loans decreases. In both cases the expected payoff from monitoring over non-monitoring shrinks. Hence, the compensation for performing monitoring activities must rise, which

intermediating mutual fund. It requires that the expected return accruing to investors - composed of the expected return from successful projects as well as the liquidation value of unsuccessful ones - is sufficient to pay back the intra-period deposits lent to the financial sector at the beginning of the period. The balance sheet constraint (3) ensures that banks' internal and external funds cover their expected expenses consisting of loans inclusive of monitoring costs related to surviving projects as well as the insurance set aside to accommodate anticipated funding outflows. Finally, (4) states that the returns accruing to individual agents add up to the total return from a successful project.

Unconstrained problem. Since the objective function is linear in project scale l, a finite solution for l can only exist when the balance sheet constraint binds. Moreover, banks maximise the amount of external resources by demanding the smallest feasible compensation for themselves and paying the smallest amount to depositors that ensures their participation. All constraints will hence bind at the optimum. Combining binding constraints (2) and (3) yields the loan scale as a function of bank capital and the liquidity cut-off:

$$l = \frac{a}{H(\bar{\omega})} \tag{5}$$

where $H(\bar{\omega}) \equiv 1 + F(\bar{\omega})\mu - q\left(1 - \int_0^{\bar{\omega}} \omega \,\mathrm{d}F(\omega)\right)\left(F(\bar{\omega})\pi_H R^h + (1 - F(\bar{\omega}))\xi^*\right)$. Banks' loan scale is thus linear in their capital with a leverage ratio of $H(\bar{\omega})^{-1}$. Plugging the loan function back into the objective function yields the unconstrained problem with the liquidity threshold $\bar{\omega}$ left as the only choice variable:

$$\max_{\{\bar{\omega}\}} \frac{F(\bar{\omega})}{H(\bar{\omega})} q\pi_H R^b a \tag{6}$$

Indifference threshold. Note that similar to the upper bound for the liquidity threshold $\bar{\omega}_1$, there is a natural lower bound $\bar{\omega}_0$. Suppose banks want to maximize the scale of their loan project. As bank capital is accumulated from retained earnings and therefore fixed, this is equivalent to attracting the largest possible amount of external financing, which consists of the sum of initial funding and the refinancing of liquidity shortfalls.[13]

Intuitively, the outside investor, i.e. the mutual fund, is willing to provide external financing until the benefit of continuation of a loan project equals its liquidation value. Conveniently, maximising the total amount of external finance, i.e. the initial loan financing plus subsequent refinancing, can be reduced to the choice of the liquidity threshold. In particular, the loan scale is pinned down as a function of banks' liquidity buffer by the participation and balance sheet

reduces the amount of loan returns that is pledgeable to outside investors.

[13]Since capital is fixed at the time of contracting, maximizing external funding is tantamount to achieving the highest feasible leverage.

constraints, (2) and (3). Maximizing external funding thus amounts to choosing the threshold for liquidity reserves $\bar{\omega}_0$ that maximizes the loan scale according to the funding constraint (5), i.e. $\max_{\bar{\omega}} \frac{a}{H(\bar{\omega})} = \frac{a}{H(\bar{\omega}_0)}$. To interpret this result, consider the corresponding first order condition, which can be expressed as[14]

$$\underbrace{\underbrace{q\pi_H R^h l}_{\text{pledgeable return}} - \underbrace{\left(1-\int_0^{\bar{\omega}_0} \omega \, \mathrm{d}F(\omega)\right)^{-1} (\bar{\omega}_0 d + \mu l)}_{\text{continuation cost}}}_{\text{continuation value}} = \underbrace{q\xi^* l}_{\text{liquidation value}}$$

Indeed, this condition suggests that at $\bar{\omega}_0$, the minimal continuation value of a loan project from the perspective of the outside investor equals its liquidation value. The minimal continuation value is the difference between the pledgeable return from the project and its maximum continuation cost. This cost consists of the highest amount $\bar{\omega}_0 d$ that may be refinanced and monitoring costs μl. Both are scaled up by the fraction of deposits that is retained by the mutual fund needing to bear these costs.

Intuitively, for any liquidity threshold $\bar{\omega} < \bar{\omega}_0$, the fund always prefers continuation of the loan project because the return even after refinancing liquidity shocks up to $\bar{\omega} d$ exceeds the project's liquidation value. The mutual fund would, in fact, be willing to expand its funding of liquidity shortfalls until being indifferent between continuation and liquidation of a project. Anticipating this incentive at the contracting stage, banks will never choose any buffer below the indifference threshold in order to take full advantage of external funding.

Optimal liquidity threshold. Since external resources are maximized at the indifference threshold for liquidity reserves, choosing liquidity reserves in excess of this threshold inevitably reduces the amount of total available external funding.[15] Thus, for any $\bar{\omega}_0 < \bar{\omega} < \bar{\omega}_1$ a trade-off emerges between total external funding and liquidity reserves. Given a fixed amount of bank capital, lower external financing will also tighten the funding constraint (5) and reduce the loan scale.

Despite this trade-off, banks optimally choose liquidity reserves in excess of the indifference threshold. To see this, note that - besides the loan scale - banks' expected return also depends on the survival probability of loan projects $F(\bar{\omega})$, which increases monotonically in the choice of the liquidity threshold $\bar{\omega}$. Therefore, banks will not seek to maximize their loan scale by setting $\bar{\omega} = \bar{\omega}_0$, but rather choose a liquidity buffer that lies between the indifference and the first-best threshold. These results are formally stated in

Proposition 1:

The liquidity threshold $\bar{\omega}$

[14] See appendix A.1 for a derivation.

[15] Analytically, this can be seen by considering the partial derivative of the leverage ratio, $\left.\frac{\partial H(\bar{\omega})^{-1}}{\partial \bar{\omega}}\right|_{\bar{\omega}>\bar{\omega}_0} < 0$. See appendix A.1 for a derivation.

(i) lies in the interval $(\bar{\omega}_0, \bar{\omega}_1)$, where the lower bound $\bar{\omega}_0$ designates the indifference threshold, which is implicitly determined by

$$\bar{\omega}_0 = \frac{\left(1 - \int_0^{\bar{\omega}_0} \omega \, \mathrm{d}F(\omega)\right)\left(\pi_H R^h - \xi^*\right) - \frac{\mu}{q}}{F(\bar{\omega}_0)\pi_H R^h + (1 - F(\bar{\omega}_0))\xi^*} \tag{7}$$

(ii) the upper bound $\bar{\omega}_1$ is the first-best liquidity threshold, which is implicitly determined by

$$\bar{\omega}_1 = \frac{\pi_H R - \xi^*}{F(\bar{\omega}_1)\pi_H R + (1 - F(\bar{\omega}_1))\xi^*} \tag{8}$$

(iii) and its optimal value is implicitly determined by the first-order condition of problem 6

$$1 = q\left[\left(1 - \int_0^{\bar{\omega}} \omega \, dF(\omega)\right)\xi^* + \bar{\omega}F(\bar{\omega})\left(F(\bar{\omega})\pi_H R^h + (1 - F(\bar{\omega}))\,\xi^*\right)\right] \equiv Q(\bar{\omega}) \tag{9}$$

where $\xi^* = \xi + z^\xi$.

Proof. See appendix A.1.

Comparative statics. The tension between liquidity reserves and loan scale is the key mechanism that links financial sector outcomes with macroeconomic dynamics. The optimal liquidity buffer fluctuates with the exogenous or pre-determined factors entering the financial contract, and, in turn, affects the amount of lending to entrepreneurs and capital production in the economy. To develop an intuition for the simulation results in section 3, consider the impact of changes in the stochastic liquidation value of loan projects ξ^* and the capital price q on optimal liquidity reserves. Under very mild parameter restrictions, the liquidity threshold correlates negatively with the liquidation value[16]

$$\frac{\partial \bar{\omega}}{\partial \xi^*} = -\frac{\partial Q / \partial \xi^*}{\partial Q / \partial \bar{\omega}} < 0 \tag{10}$$

Intuitively, a lower liquidation value reduces the incentive of the outside investor to terminate a project after liquidity withdrawals have occurred, such that the indifference threshold $\bar{\omega}_0$ increases. Thus, banks are able to extract more refinancing from the mutual fund *ex post* at any initial loan scale. In other words, the participation constraint of the outside investor becomes less sensitive to the liquidity threshold when the liquidation value shrinks. This shifts the *ex ante* trade-off between liquidity reserves and the loan scale in favour of a higher accumulation of reserves from the perspective of banks. Note that the trade-off does not disappear altogether, such that a higher liquidity threshold will still constrain the initial lending scale. However, with a lower liquidation value the contraction in the lending scale will be less severe for a given increase in the liquidity

[16] The partial derivatives of $\bar{\omega}$ are derived in appendix A.2.

buffer.

The optimal liquidity threshold is also negatively correlated with the capital price, i.e.

$$\frac{\partial \bar{\omega}}{\partial q} = -\frac{\partial Q/\partial q}{\partial Q/\partial \bar{\omega}} < 0 \tag{11}$$

A lower asset price q decreases the market value of loan projects, and thus their marginal profitability. When the profitability of loan projects is low the opportunity cost of shifting resources to the liquidity buffer is lower. In other words, the balance-sheet trade-off between scale and reserves is weakened when the asset price falls. As a result, banks increase their liquidity reserves, which increases the survival probability of loans.

2.3.2 The Mutual Fund and Endogenous Liquidity Supply

The mutual fund intermediates households' deposits at the beginning of subperiod 2 to provide external funding to banks. When loan contracts have been completed at the end of subperiod 4, it collects the pledgeable proceeds of loans from banks and channels them to depositors. The fund finances initial loans and also acts as a liquidity backstop for the banking sector by partially insuring liquidity risks. Although deposit withdrawals are idiosyncratic at the bank-level, the mutual fund cannot diversify them up to the first-best threshold $\bar{\omega}_1$, as discussed in section 2.3, because of the moral hazard problem faced by banks.

The insurance scheme allows banks to draw on a contingent credit line only up to a maximum amount $\bar{\omega}d < \bar{\omega}_1 d$ after funding outflows, which is compatible with their incentive to provide monitoring services. However, banks do not set aside the full amount as a buffer with the mutual fund, but rather an amount equal to the expected liquidity shock $\left(\int_0^{\bar{\omega}} \omega \, \mathrm{d}F(\omega)\right) d < \bar{\omega}d$. To understand how the mutual fund can refinance liquidity outflows up to the promised amount, note that liquidity shocks are independent across banks by assumption. Since there is a continuum of banks, the expectation of the refundable funding outflow from any bank is equal to the aggregate funds actually withdrawn from the financial sector *ex post*. Therefore, aggregate liquidity demand that the fund is asked to refinance in subperiod 3 is deterministic and given by

$$W = \left(\int_0^{\bar{\omega}} \omega \, \mathrm{d}F(\omega)\right) D$$

where $D = \eta^b d$. In order to provide this amount, the mutual fund redistributes banks' liquidity reserves from those with low outflows to those with high outflows. However, in order to ensure the participation of depositors, the fund can only satisfy the aggregate liquidity demand as long as it does not exceed the pledgeable returns to loans. The latter are equal to the market value of the banking sector in subperiod 3, which amounts to

$$\begin{aligned} V &= q\left\{F(\bar{\omega})\pi_H R^h + (1 - F(\bar{\omega}))\xi^*\right\} L \\ &= D \end{aligned}$$

where $L = \eta^b l$. This market value always exceeds banks' liquidity need as the fraction of liquidity outflows is bounded by 1 from above, such that $V - W = \left[1 - \int_0^{\bar{\omega}} \omega \, \mathrm{d}F(\omega)\right] D \geq 0$.[17] Hence, there is no aggregate shortage of valuable claims on the banking sector when liquidity shocks arrive, such that the insurance scheme up to the optimal liquidity threshold $\bar{\omega}$ is feasible.[18]

Since the aggregate refundable liquidity demand in the banking sector is deterministic, the fund can offer a riskless rate of return to depositors, which ensures risk neutrality of households with respect to deposits.

2.3.3 Evolution of Bank Capital

Each period, $1 - \tau^b$ bankers exit the financial sector and are replaced by a continuum of new bankers of the same mass. The share of bankers in the economy thus stays constant at η^b. Bankers save the proceeds from their intermediation activity by accumulating capital k_t^b. They derive income from renting their capital out to final goods producers and supplying one unit of labour inelastically to the same sector. After final goods production is completed, they earn the respective factor rents. Labour income provides small positive start-up funds even to assetless new bankers. Bank capital in subperiod 2 thus equals

$$a_t = (q_t(1 - \delta) + r_t)k_t^b + w_t^b \tag{12}$$

Each banker invests his entire capital into a loan project yielding $R_t^b l_t$ if successful and zero otherwise. The proceeds can either be saved or consumed. The inter-temporal flow of funds of individual banks is

$$\begin{aligned} c_t^b + q_t k_{t+1}^b &= (1 + r_t^a)a_t \\ &= q_t F(\bar{\omega}_t)\pi_H R_t^b l_t \end{aligned}$$

where $1 + r_t^a = \frac{q_t F(\bar{\omega}_t)\pi_H R_t^b}{H(\bar{\omega})}$ is the gross return on bank capital and the last line uses equation (5). Successful surviving bankers save the entire proceeds from their lending activity in capital goods.

[17] In particular, given a uniform distribution of the liquidity shock on the interval $[0, U_H]$, where $U_H \in [0, 2]$:

$$\bar{\omega} < U_H < \sqrt{2U_H} \quad \Longrightarrow \quad \int_0^{\bar{\omega}} \omega \, \mathrm{d}F(\omega) < 1 \quad \Longrightarrow \quad V - W > 0$$

[18] Other arrangements that achieve the same risk-sharing outcome, such as banks directly holding a stake in the market portfolio of the banking sector, are discussed in Holmström and Tirole (2011).

This is the optimal consumption-savings choice given bankers' risk-neutrality and the high return on internal funds.[19] Bankers' whose projects yield no return lose all their capital and, accordingly, neither save nor consume. Exiting bankers consume their entire assets.

The ad hoc assumption of a finite lifetime for bankers ensures the stationarity of aggregate bank capital. If bankers did not exit the economy to consume their assets they would eventually accumulate enough wealth to finance investments exclusively with internal funds.[20]

2.4 Final Good Producers

Final good producers operate on a competitive, frictionless market. They use the aggregate capital stock K_t rented from households and bankers and aggregate labour supplied by households H_t^h and bankers H_t^b as inputs into production.

$$Y_t = \exp(z_t)F(K_t, H_t^h, H_t^b) \tag{13}$$

where $\exp(z_t)$ is total factor productivity. Factors earn their marginal product, such that the interest rate on capital is $r_t = \exp(z_t)F_K(K_t, H_t^h, H_t^b)$ and wages are given by $w_t^i = \exp(z_t)F_{H^i}(K_t, H_t^h, H_t^b)$ for $i \in \{b, h\}$.

2.5 Households

There exists a continuum of households of mass η^h. Households are risk averse and maximise utility over consumption c_t^h and labour h_t^h subject to their individual budget constraints. At the beginning of each period, households lend previously accumulated capital k_t^h to final goods producers and supply labour to the same sector. Both factors are remunerated with their respective rents. Likewise, last period's bonds pay a gross riskless return $1+r_t^b$. Capital depreciates at rate δ. Then households make their consumption-savings decision. In order to save, they have two options: Purchasing one-period risk-free bonds or channelling funds to banks via the mutual fund. After banks have performed their intermediation activity and investment projects generate returns, q_t units of new capital goods are transferred to households for every unit of savings input. Choosing the amount of deposits is thus equivalent to choosing how much capital to hold in the future.

[19] The model calibration ensures that the marginal benefit always exceeds the marginal cost of saving for surviving bankers, i.e. $\left(1 + r_{t+1}^a\right)(q_{t+1}(1-\delta) + r_{t+1}) > q_t$.

[20] This is a well-known property of macroeconomic models with financially constrained agents, shared, for instance, by Bernanke, Gertler, and Gilchrist (1999), Gertler and Karadi (2011), Gertler and Kiyotaki (2011) or Christiano, Motto, and Rostagno (2014).

Accordingly, the optimization problem takes the form

$$
\begin{aligned}
\max_{\{c_t^h, k_{t+1}^h, b_{t+1}^h, h_t^h\}} \quad & E_0 \sum_{t=0}^{\infty} \beta^t u(c_t^h, h_t^h) \\
& s.t. \\
& c_t^h + q_t k_{t+1}^h + b_{t+1}^h = (1 + r_t^b) b_t^h + (q_t(1-\delta) + r_t) k_t^h + w_t^h h_t^h
\end{aligned} \tag{14}
$$

The corresponding first order conditions for consumption, capital stock, bonds and labour supply read

$$
\begin{aligned}
u_{c,t} &= \lambda_t & (15) \\
\lambda_t &= \beta E_t \left[\lambda_{t+1} \frac{q_{t+1}(1-\delta) + r_{t+1}}{q_t} \right] & (16) \\
\lambda_t &= \beta E_t \left[\lambda_{t+1} (1 + r_{t+1}^b) \right] & (17) \\
u_{h,t} &= -\lambda_t w_t^h & (18)
\end{aligned}
$$

where (16) and (17) are the Euler equations with respect to capital and bonds, respectively.

2.6 Aggregation and Competitive Equilibrium

Due to linearities in the financing and production of capital goods, aggregation turns out to be straightforward. In particular, the production technology for new capital goods and monitoring costs are linear in loans. The distribution of bank capital, therefore, has no effect on aggregate loans L_t and investment $I_t = L_t$, which are simply the sum of individual loans:

$$
\begin{aligned}
L_t &= \eta^b l_t \\
&= \frac{\eta^b a_t}{H(\bar{\omega}_t)} = \frac{A_t}{H(\bar{\omega}_t)}
\end{aligned} \tag{19}
$$

using the individual loan function (5).

The economy-wide equivalent to depositors' participation constraint (2) pins down aggregate deposits.

$$
\begin{aligned}
D_t &= \eta^b d_t \\
&= q_t \left\{ F(\bar{\omega}_t) \pi_H R_t^h + (1 - F(\bar{\omega}_t)) \xi^* \right\} L_t
\end{aligned} \tag{20}
$$

Aggregate stocks of capital holdings are the sum of individual stocks.

$$
K_t^b = \eta^b k_t^b, \quad K_t^h = \eta^h k_t^h \tag{21}
$$

The elasticity of labour supply differs across agents. Bankers individually supply one unit of labour inelastically, while households' supply is elastic.

$$H_t^b = \eta^b, \quad H_t^h = \eta^h h_t^h \tag{22}$$

Aggregate bank capital is

$$A_t = (q_t(1-\delta) + r_t)K_t^b + H_t^b w_t^b \tag{23}$$

The average return on loans for bankers is $F(\bar{\omega})\pi_H R^b l_t$. As discussed, surviving bankers invest all their proceeds into new capital goods. Since only a fraction τ^b survives, next period's capital holdings by the banking sector will be

$$K_{t+1}^b = \tau^b F(\bar{\omega}_t)\pi_H R_t^b L_t \tag{24}$$

Exiting bankers consume their wealth and aggregate household consumption amounts to the sum of individual households' consumption.

$$C_t^b = (1-\tau^b)q_t F(\bar{\omega}_t)\pi_H R_t^b L_t \tag{25}$$

$$C_t^h = \eta^h c_t^h \tag{26}$$

The competitive equilibrium of the economy is a collection of (i) decision rules for $c_t^h, k_{t+1}^h, b_{t+1}^h, h_t^h$ that solve the maximization problem of households; (ii) decision rules for K_t, H_t^b, H_t^h that solve the maximization problem of final good producers; (iii) decision rules for $l_t, d_t, R_t^b, R_t^h, \bar{\omega}_t$ associated with the financial contract that solves the maximization problem of banks; (iv) consumption c_t^b and saving k_{t+1}^b rules for bankers; (v) laws of motion for the exogenous processes z_t, z_t^ξ, and market clearing conditions for final goods, labour, capital goods, investment, loans and bonds[21] :

$$C_t^h + q_t K_{t+1}^h + B_{t+1} = (1+r_t^b)B_t + (q_t(1-\delta)+r_t)K_t^h + w_t^h H_t^h \tag{27}$$

$$H_t = H_t^b + H_t^h \tag{28}$$

$$K_t = K_t^b + K_t^h \tag{29}$$

$$K_{t+1} = (1-\delta)K_t + \left(F(\bar{\omega}_t)\pi_H R + (1-F(\bar{\omega}_t))\xi\right) I_t \tag{30}$$

$$q_t L_t = q_t I_t \tag{31}$$

$$B_t = 0 \tag{32}$$

[21] Appendix A.3 lists the complete set of equilibrium conditions.

3 Quantitative Results

3.1 Calibration and Functional Forms

Period-utility - a function of consumption and hours worked - takes the following functional form:

$$u(c_t^h, h_t^h) = \frac{c_t^{h^{1-\theta}}}{1-\theta} + \nu \ln(1 - h_t^h) \tag{33}$$

The parameter θ governs the degree of relative risk aversion or the elasticity of intertemporal substitution of consumption. It is set to a standard value of 1.5 following Kato (2006). The weight on leisure, ν, is chosen to match a fraction of working time of 30%. Additionally, households' discount factor is set to a standard value of 0.99, which yields a riskless quarterly interest rate of 1%.

Final goods are produced with a standard Cobb-Douglas technology

$$F(K_t, H_t^h, H_t^b) = K_t^{\alpha^k} H_t^{h^{\alpha^h}} H_t^{b^{\alpha^b}} \tag{34}$$

where $\alpha^k + \alpha^h + \alpha^b = 1$. I follow Meh and Moran (2010) in setting the capital share of output to 0.36 and the share of labour provided by bankers to a very small number (5×10^{-5}), such that its effect on the dynamics is negligible.

Capital production is characterized by two parameters. A quarterly depreciation rate of capital of $\delta = 0.025$ is in line with many RBC studies of the US economy including King and Rebelo (1999), Kato (2006) and Covas and Fujita (2010). There is less precedent for the second parameter choice, R, i.e. the return to investment in capital production. I calibrate this parameter such that the total return to investment with full buffering of liquidity shocks is one, i.e. $\pi_H R = 1$.[22]

Financial intermediation and the associated frictions are characterized by the set of parameters $\{\mu, \xi, \sigma^2(\omega), \tau^b, \pi_H, \pi_L\}$. The parameters π_H and π_L capture the idiosyncratic failure risk of entrepreneurs under effort and shirking. Following Meh and Moran (2010), I set $\pi_H = 0.9903$, which translates into a quarterly failure rate of entrepreneurs of 0.97%, as in Carlstrom and Fuerst (1997), and $\pi_L = 0.75$.

The subset $\{\mu, \xi, \sigma^2(\omega), \tau^b\}$ is jointly determined to match: (i) A bank-leverage ratio, defined as the ratio of debt to equity $\Xi_t = \frac{D_t}{A_t}$, close to 13.44. This roughly corresponds to the average leverage ratio of the US financial sector composed of banks and market-based financial institutions over the past 30 years (Figure 6).[23] (ii) A loss given default (LGD) on bank loans of roughly 40% following

[22] In this case, the agency cost model collapses to the standard real business cycle model as consumption goods are converted into capital goods one-to-one *ex post*.

[23] For issuers of asset-backed securities (ABS), which make up an important fraction of the market-based intermediation sector as demonstrated in Figure 7, no data on leverage ratios was available. Since market-based intermediaries' leverage tends to exceed that of traditional banks, the average leverage of 13.44 computed for the financial sector without ABS issuers is likely downward-biased. Hence, I allow for a slightly higher leverage ratio in the model.

Table 1: Baseline calibration

	Parameter	Value	Target/Source
Preferences			
Household discount factor	β	0.99	riskless interest rate: 1%
Relative Risk aversion	θ	1.50	Kato (2006)
Utility weight on leisure	ν	2.67	working time: 30%
Final goods production			
Capital share of output	α^k	0.36	Meh and Moran (2010)
Labour share of output (households)	α^h	0.63995	Meh and Moran (2010)
Labour share of output (bankers)	α^h	0.00005	Meh and Moran (2010)
Capital goods production			
Depreciation rate of capital	δ	0.025	Kato (2006)
Return to investment	R	1.0098	one-to-one transformation
Financial Intermediation			
Unit-monitoring cost	μ	0.1308	investment-to-GDP: 0.15
Liquidtation value to outsiders	ξ	0.2400	loss-given-default: 0.4
Probability of success: effort	π_H	0.9903	quarterly failure rate: 0.0097
Probability of success: shirking	π_L	0.7500	Meh and Moran (2010)
Std. dev., idiosync. liquidity risk	σ_ω	0.3458	liquidity share: 0.21
Population parameters			
Mass of households	η^h	0.97	Meh and Moran (2010)
Mass of bankers	η^b	0.03	Meh and Moran (2010)
Share of surviving bankers	τ^b	0.26	bank-leverage ratio: 15
Shock processes			
Persistence, productivity shock	ρ_a	0.90	Kato (2006)
Std. dev., productivity shock	σ_a	0.01	1%
Persistence, liquidity shock	ρ_ξ	0.90	
Std. dev., liquidity shock	σ_ξ	0.048	20%

Notes: The model is calibrated for quarterly data.

Covas and Fujita (2010). In the model, the LGD corresponds to $LGD_t = 1 - \frac{q_t \xi_t^* I_t}{I_t - A_t}$. (iii) The share of liquid assets in banks' balance sheets, $\Omega_t = \frac{\int_0^{\bar{\omega}_t} \omega \, dF(\omega) D_t}{A_t + D_t}$. As the empirical counter-part I use the sum of cash, central bank reserves as well as all government-backed assets relative to balance sheet size. The evolution of this liquidity share for banks and market-based intermediaries was shown in Figure 1. While the ratio varied widely, between 13% to 30% for banks and 2% to 23% for shadow banks, over the past three decades, the model targets the average empirical liquidity share of about 19%. (iv) An investment-to-GDP ratio $\Phi_t = \frac{I_t^{eff}}{Y_t}$ of 15%, where effective investment is defined as $I_t^{eff} \equiv [F(\bar{\omega}_t)\pi_H R_t + (1 - F(\bar{\omega}_t))\, \xi_t^*]\, I_t$.

Liquidity shocks are assumed to be distributed uniformly on the interval $[0, U_H]$, such that $\sigma^2(\omega) = \frac{U_H^2}{12}$. The assumption of a uniform distribution facilitates the analysis, but results do not depend on it. With the chosen calibration, we have $\bar{\omega}_0 = 0.68$, $\bar{\omega} = 0.73$ and $\bar{\omega}_1 = 1.04$. As the optimal liquidity threshold falls between the indifference threshold and the frictionless first-best refinancing threshold, the trade-off between the lending scale and the liquidity buffer is operative as discussed in section 2.3.1.

The full set of calibrated parameters including the remaining population parameters is listed in Table 1. Some key matched moments and their model-equivalents are summarized in Table 2.

Table 2: Selected Targets: Data vs. Model

Target	Concept	Data	Model
Loss given default	LGD	39.8%	0.40
Leverage ratio	Ξ	13.44	15
Liquidity share	Ω	18.78%	0.21

Notes: The average leverage ratio of the US financial industry is an asset-weighted average of the average leverage of bank- and market-based institutions. Due to lack of data for ABS issuers, this value is likely to be downward-biased. Data on loss given default derives from Araten et al. (2004), who report loan losses of a large US bank between 1982 and 1999. The empirical counterpart to the liquidity share is computed as the sum of checkable deposits and currency, cash and reserves at the Federal Reserve, Treasury securities and agency- and GSE-backed securities relative to total assets of the respective institutions. Sources: US Flow of Funds (Federal Reserve), Araten et al. (2004)

3.2 Aggregate shocks

I consider two types of aggregate risk in the economy. The first is a standard technology shock that follows the process $z_t = \rho z_{t-1} + e_t$.

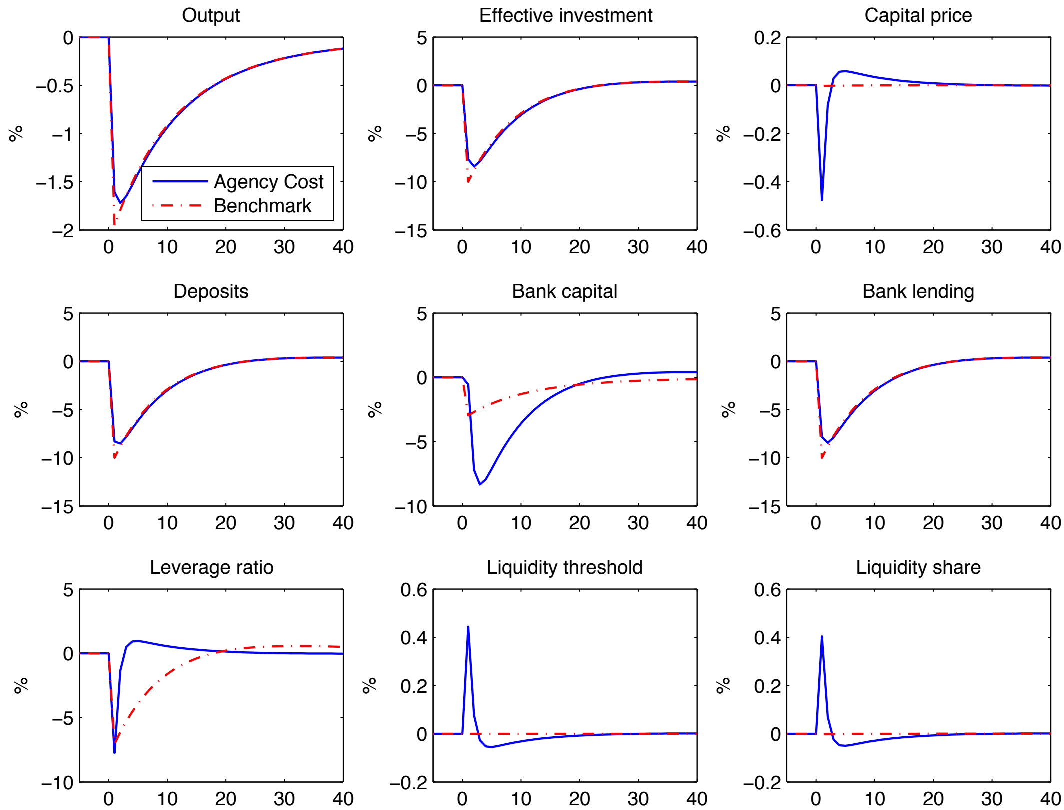
The second source of aggregate risk is a collateral shock. I model this shock as a collapse in the liquidation value of bank loans to outside investors, i.e. $\xi_t^* = \xi - z_t^\xi$ where $z_t^\xi = \rho_\xi z_{t-1}^\xi + e_t^\xi$ and $e_t^\xi \sim \mathcal{N}(0, 0.048)$. This shock to the liquidation value of loans is intended to capture the sudden decline in the collateral value of bank assets at the onset of the financial crisis. Gorton and Metrick (2010, 2012) investigate the development of the collateral value of bank assets during the Great Recession by analysing the repo market, a primary source of short-term refinancing among market-based financial intermediaries. The authors argue that haircuts on the underlying assets in repo transactions amount to a reduction in the collateral values of these assets. During the financial crisis, and particularly in the wake of the Lehman crash in September 2008, haircuts in repo contracts surged from close to zero to 45% on average. Even non-subprime-related assets suffered haircuts of up to 20%. In order to evaluate whether such a collateral shock may have contributed to the severe recession, I hit the steady-state collateral value with a (conservative) negative shock of 20%.

The following sections present my main findings regarding business cycle dynamics in the presence of idiosyncratic liquidity risk and a balance sheet channel of shock transmission working through the banking sector. The model is solved using a first-order approximation to the policy functions around the non-stochastic steady state.

3.2.1 Aggregate Technology Shock

The impulse response functions of key aggregate variables to a one-standard deviation technology shock are shown in Figure 3 along with impulse responses of the frictionless benchmark model without agency costs ($\mu = 0$), but identical technological constraints. In the absence of agency costs the first-best refinancing threshold is constant at $\bar{\omega}_1$ and does not contribute to shock amplification and transmission. As a result, impulse responses in the benchmark model resemble those of a frictionless real business cycle model, while the agency-cost model exhibits slightly hump-shaped responses, particularly in output and investment.

Figure 3: Responses to a Technology Shock



Notes: Impulse responses to a negative one percent technology shock. The agency-cost model (solid lines) is contrasted with a frictionless benchmark model where $\mu = 0$ (dashed lines).

To understand the dynamics in the agency cost model, consider that lower productivity in the final goods sector reduces factor rents. Households react to lower expected rental income from holding capital by reducing their demand for new capital goods. This, in turn, puts downward pressure on the price of capital, q_t. The fall in the capital price reduces the profitability of the banking sector, and thus the returns on deposits as can be seen from households' aggregate participation constraint (20). As a consequence, they provide less deposits to the banking sector.

At the same time, the low return on capital and the fall in the capital price reduce banks' capital contemporaneously, as suggested by equation (23). In fact, the fall in bank capital is much more pronounced than in the frictionless benchmark model due to the drop in the value of capital goods. With both external and internal funding eroding, banks are forced to shorten their balance sheets by curtailing their lending.

Further pressure on the volume of bank loans comes from the response of the liquidity buffer. As discussed in section 2.3.1, the liquidity threshold chosen by banks is negatively correlated with the relative price of capital, because the opportunity cost of holding liquidity reserves is lower when the capital price and the profitability of loans fall. Hence, $\bar{\omega}$ increases slightly. The share of liquid assets in banks' balance sheets, Ω_t, increases mainly on account of the balance sheet contraction (denominator effect). Overall, financial adjustments related to liquidity hoarding are rather short-lived.

Output drops immediately due to lower factor productivity. Its hump-shaped response derives from the sluggishness of bank capital. The latter continues to drop after the initial shock because of the decrease in the capital price. Lower bank capital propagates via more sluggish lending activity into capital formation and future output.

3.2.2 Aggregate Collateral Shock and the Great Recession

As argued in section 2.3.1, a negative shock to the collateral value of bank assets shifts the trade-off between loan scale and liquidity reserves in favour of the latter. The collateral shock thus unfolds its effects through the increase in banks' liquidity buffers and the corresponding reduction in lending. These dynamics can be traced in the impulse responses shown in Figure 4.

In order to disentangle the effect of agency problems and liquidity hoarding from the impact of the collateral shock as such consider first the frictionless benchmark version of the model. In the absence of agency problems between investors and banks liquidity shocks are insured up to the first-best threshold $\bar{\omega}_1$, such that no projects whose continuation value exceeds the liquidation value are abandoned. Since the refinancing threshold is constant, the collateral shock only affects the economy through its direct impact on the participation constraint of the mutual fund. In particular, the fund reduces its overall funding to the banking sector in view of the lower expected liquidation value of loan projects. Given the pre-determined nature of banks' capital at the time of contracting, banks are forced to react to the drop in external financing by curtailing lending. Recall the definition of effective investment as

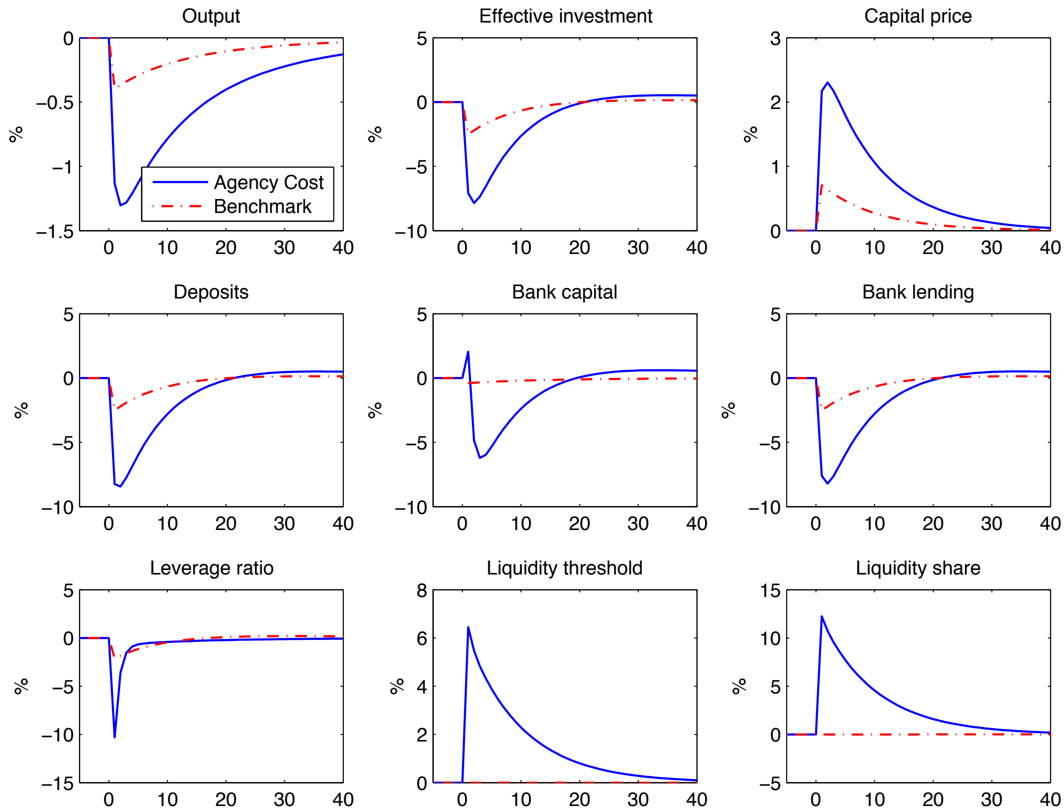
$$I_t^{eff} = \left(F(\bar{\omega}_1)\pi_H R + (1 - F(\bar{\omega}_1))\xi_t^*\right) I_t \tag{35}$$

Investment falls both directly due to the impact of the lower liquidation value of failed investment projects, as well as indirectly through the reduction of the loan, and thus investment scale. The drop in effective investment reduces capital accumulation, which propagates the shock into the future.[24]

In the economy perturbed by agency conflicts, banks respond to a drop in the collateral value ξ_t^* by increasing their liquidity threshold $\bar{\omega}_t$ as implied by the first-order condition (9). The higher liquidity threshold raises the survival probability of investment projects (marginal effect), which would of itself increase effective investment. However, given the trade-off between liquidity reserves and loan scale, higher liquidity reserves come at the cost of an amplified contraction in lending

[24] The increase in the price of capital is closely linked to the investment dynamics: The collateral shock essentially triggers a negative supply shock on the capital market. This is not matched by a decline in demand for capital, since capital is still highly productive. Therefore, the capital price increases, which in turn boosts households' net worth and, hence, consumption. This shortcoming is shared by models with shocks directly affecting investment or capital supply and without alternative means of saving for households (Del Negro, Eggertsson, Ferrero, and Kiyotaki, 2011; Kiyotaki and Moore, 2012).

Figure 4: Responses to a Collateral Shock



Notes: Impulse responses to a negative 20% percent collateral shock. The agency-cost model (solid lines) is contrasted with a frictionless benchmark model where $\mu = 0$ (dashed lines).

when compared with the frictionless economy (scale effect). The strong fall in effective investment suggests that the negative shock to the liquidation value in combination with the contraction in lending clearly dominates the effect of higher liquidity reserves on effective investment. Although the relative importance of the marginal versus the scale effect is an empirical question, intuitively the scale effect should be expected to dominate as it works through banks' leverage (see equation (19)), which is highly sensitive to changes in the liquidity threshold.[25]

Since the collateral shock directly affects the choice between banks' lending scale and liquidity reserves, the response of the liquidity threshold $\bar{\omega}_t$ follows the shock's AR(1) structure and is, thus, much more prolonged than in the case of a technology shock. The sustained increase in the liquidity share Ω_t reflects both the higher demand for liquidity reserves (numerator effect) and the contraction in banks' balance sheets (denominator effect). Contrasting these results with the response of the frictionless benchmark economy, where the trade-off between liquidity and scale is

[25] Recall from equation (19), that the loan scale has a highly non-linear relationship with the liquidity threshold $\bar{\omega}$ through the leverage ratio $H(\bar{\omega})^{-1}$ and, therefore, reacts very sensitively to changes in $\bar{\omega}$ (see proof A.1). In comparison, the survival probability of loan projects, i.e. the cumulative distribution function of $\bar{\omega}$, is much less sensitive to changes in the liquidity threshold by comparison (in fact, it is linear in the parameterization at hand), such that the leverage - or balance sheet - contraction dominates.

absent, reveals *liquidity hoarding* by financial intermediaries as the key amplification mechanism of the initial collateral shock.

Propagation works in much the same way as in the benchmark model. Depressed investment eats into banks' capital stock, forcing them to curtail lending in future periods as well. The sluggish response of bank capital thus translates into hump-shaped lending, investment, and output. Interestingly, the model is able to replicate this hump-shaped response without recourse to adjustment costs, solely through balance sheet dynamics. As a second consequence of the sluggishness of bank capital relative to deposits, bank leverage becomes pro-cyclical. The model can thus rationalize both the scramble for liquidity and the strong deleveraging of financial intermediaries observed in the data during the Great Recession (Figures 1 and 6, respectively).[26]

Although the initial aggregate shock is amplified through a balance sheet channel, the effects in this model are quite distinct from the financial accelerator framework. In that framework, fluctuations in borrower net worth affect the borrowing capacity of financially constrained agents. Negative shocks to borrowers' net worth induce fire-sales which increase the initial losses and lead to further fire-sales. In the present model, the amplification mechanism works instead through the composition of the asset side of constrained borrowers' balance sheets between liquidity buffers and loan scale. Borrowers' net worth simply drops as a consequence of the negative impact of the credit crunch on investment and the capital stock, but is not the cause of the crunch. Hence, the model develops a novel type of shock transmission through borrowers' balance sheets.[27]

The key insight from this analysis is that even a modest drop in the collateral value of assets held in the financial sector triggers a flight to liquidity associated with output losses of 1.25%. During the financial crisis, average haircuts in repo contracts were more than twice as high as those modelled in this paper. The counter-cyclical flight to liquidity channel described here may thus have been an important amplification mechanism during the Great Recession.

3.3 Financial and Nominal Frictions

During the Great Recession, financial frictions are likely to have interacted with nominal rigidities. Christiano, Trabandt, and Walentin (2011), for instance, emphasize the Fisherian debt-deflation mechanism according to which deflationary pressures inflate the real value of nominal debt. At the

[26] Note that the amplified drop in investment is tantamount to a larger negative supply shock on the capital goods market. Hence, the price reacts more strongly, increasing households' net worth and consumption. This rise of consumption on impact prevents the economy from sliding into an even deeper recession. The initial jump in bank capital is also generated by the increase in the price of capital.

[27] Financing constraints induced by agency frictions make the present model sensitive to the distribution of wealth among agents. Therefore, the amplification mechanism presented here could be attenuated by transferring resources from unconstrained households to constrained banks. Such wealth transfers in the form of capital injections or public debt guarantees were the cornerstone of unconventional crisis policies adopted, for instance, by the US Government during the course of the financial crisis. They were ultimately aimed at restoring the capital and borrowing capacity of financial intermediaries. The present model lends itself to the study of such unconventional policies, which would reduce the recessionary impact of collateral shocks. However, since a fully-fledged welfare analysis is beyond the scope of the paper I do not pursue this avenue further.

same time, nominal frictions affect the consumption-savings decisions of households through their impact on the real interest rate. For instance, rising real interest rates are key for explaining the strong output losses experienced during the Great Recession in the model of Del Negro, Eggertsson, Ferrero, and Kiyotaki (2011). A similar mechanism deepens the recession triggered by a collateral shock in the present model.

To add nominal rigidities in product markets I assume an additional layer in the production process in the form of monopolistically competitive intermediate goods producers. Final goods are assembled from intermediate goods via a standard Dixit-Stiglitz aggregation technology with finite elasticity of substitution between different varieties of intermediate goods. Intermediary producers use their market power to price their goods at a mark-up over marginal costs. Moreover, they face price adjustment costs as in Rotemberg (1982), such that they do not adjust prices fully in response to variations in demand for their respective goods. Optimal price setting, thus, yields the familiar forward-looking New Keynesian Phillips curve.[28]

$$\left(\frac{\pi_t}{\bar{\pi}}-1\right)\left(\frac{\pi_t}{\bar{\pi}}\right)=\frac{\epsilon}{\chi}\left(m_{c,t}-\frac{\epsilon-1}{\epsilon}\right)+\mathbb{E}_t\left[\frac{\beta\lambda_{t+1}}{\lambda_t}\left(\frac{\pi_{t+1}}{\bar{\pi}}-1\right)\left(\frac{\pi_{t+1}}{\bar{\pi}}\right)\frac{Y_{t+1}}{Y_t}\right] \tag{36}$$

Monetary policy is assumed to react to deviations of inflation and output from their respective non-stochastic steady states according to the following rule:

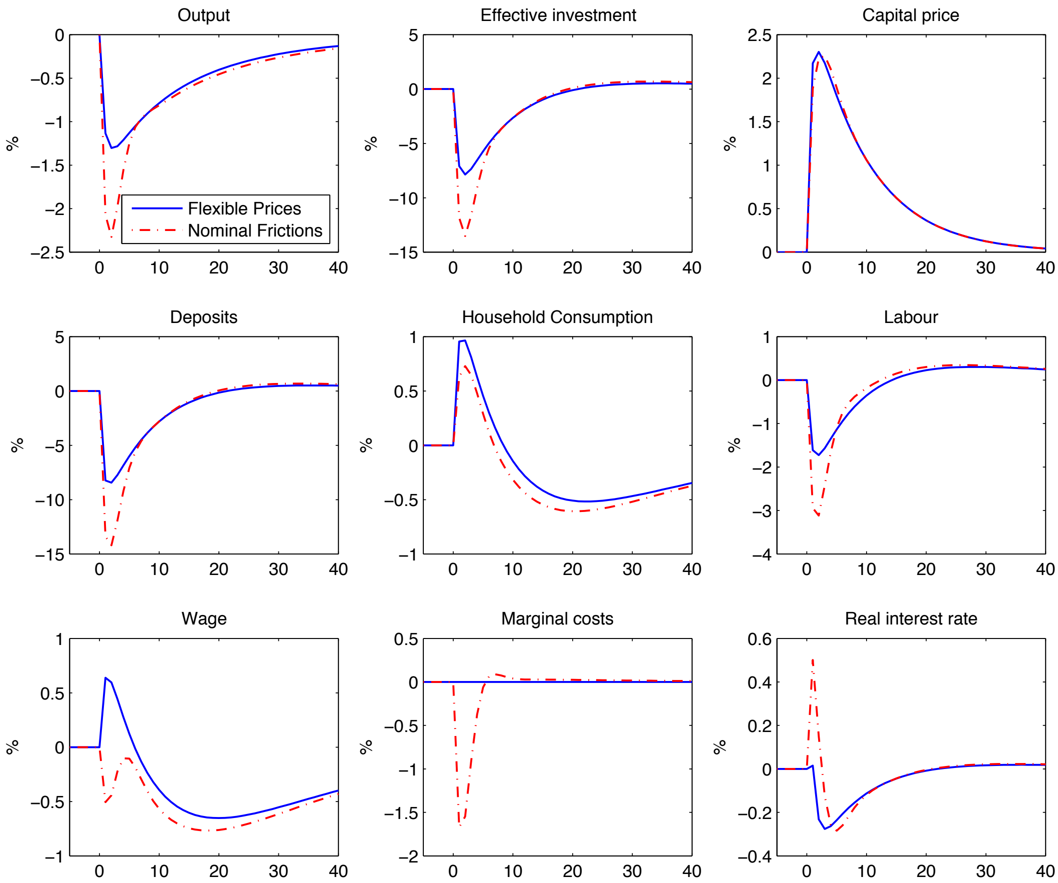
$$i_t^b=(1-\rho_i)i^b+\rho_i i_{t-1}^b+(1-\rho_i)\left[\rho_\pi(\pi_t-\bar{\pi})+\rho_y(Y_t-\bar{Y})\right] \tag{37}$$

I calibrate the elasticity of substitution between intermediate good varieties to $\epsilon=6$ and the parameter governing price adjustment costs to $\chi=29$. These choices are consistent with estimates of the slope coefficient of the log-linear Phillipps curve as derived from the Calvo-Yun model (Galí and Gertler, 1999). The coefficients of the policy reaction function derive from those estimated in Clarida, Galí, and Gertler (2000), i.e. $\rho_\pi=1.5$, $\rho_i=0.8$, $\rho_y=0.1$.

As the impulse responses in Figure 5 show, nominal frictions exacerbate the effect of a collateral shock on output significantly compared to the baseline model with flexible prices. The stronger decline in output, particularly in the first four quarters after the shock, results both from a further decline in investment as well as a more muted rise in consumption. As the increasing capital price raises households' net worth after a collateral shock in the flexible-price baseline, they increase consumption. With price rigidities, however, the collateral shock allows inflation to drop below the steady state. Since the nominal interest rate lags economic dynamics, this causes the real interest rate to rise until the monetary authority reacts by cutting the nominal rate more aggressively. *Ceteris paribus*, higher real interest rates would tilt households towards saving rather than consuming. However, marginal costs drop with inflation such that wages are marked down further and hours worked fall strongly. Hence, households' net worth declines leading to both lower

[28] For a detailed derivation see Appendix A.5.

Figure 5: Responses to a Collateral Shock with Nominal Frictions



Notes: Impulse responses to a negative 20% collateral shock with (dashed) and without nominal frictions (solid).

savings and a reduced increase in consumption. In the absence of a substantial private demand boom the output decline is much stronger. The interaction of financial with nominal frictions is, hence, a key factor in explaining the severity of the recession. This observation underscores the importance of effective monetary policy and hints at the potential distortions introduced by the zero lower bound.

4 Conclusion

This paper studies the impact of idiosyncratic funding liquidity risk in the presence of financial frictions between banks and investors on business cycle fluctuations. A standard moral hazard problem induces a skin-in-the-game constraint which forces banks to retain an equity stake in their loans. The same moral hazard problem that puts a limit on bank loans at the initial lending stage also constrains the amount of funding outflows that can be refinanced during the lifetime of loans. Idiosyncratic liquidity shocks can thus not be fully diversified and may lead to the termination of highly productive loan projects. Anticipating this risk, banks reduce their initial lending scale in order to set aside resources as liquidity buffers. Hence, balance sheet constraints force banks to trade off insurance against idiosyncratic liquidity risk with initial loan scale.

A shock to the collateral value of bank assets is introduced as a novel source of aggregate risk, which directly operates on the participation constraint of banks' outside investors. Banks react to such a shock by hoarding more liquidity at the expense of their lending scale. In the aggregate, this scale effect combined with the lower liquidation value of terminated loan projects dominates the higher survival probability of loans, such that net investment falls and economic activity contracts sharply. Decreases in bank capital propagate shocks through time and induce a hump-shaped response of output. This credit crunch scenario shares key aspects with the Great Recession, which was triggered by losses on financial assets resulting in a flight to liquidity and a lending squeeze. Furthermore, the interaction of nominal with financial frictions is shown to amplify the business cycle dynamics stemming from the flight to liquidity channel.

The models identifies a new, quantitatively important type of amplification mechanism working through endogenous portfolio choices of financial intermediaries in the presence of idiosyncratic funding liquidity risk. This paper thus contributes to the growing body of literature merging macroeconomic models with financial frictions.

References

ADRIAN, T., AND H. S. SHIN (2009): "Money, Liquidity, and Monetary Policy," *American Economic Review: Papers & Proceedings*, 99(2), 600–605.

ARATEN, M., M. JACOBS JR., AND P. VARSHNEY (2004): "Measuring LGD on Commercial Loans: An 18-Year Internal Study," *RMA Journal*, 86(8), 28–35.

BERNANKE, B., AND M. GERTLER (1989): "Agency Costs , Net Worth , and Business Fluctuations," *American Economic Review*, 79(1), 14–31.

BERNANKE, B., M. GERTLER, AND S. GILCHRIST (1999): "The Financial Accelerator in a Quantitative Business Cycle Framework," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1(3) of *NBER working papers series*, chap. 21, pp. 1341–1393. Elsevier Science B.V.

BERNANKE, B. S., AND M. GERTLER (1995): "Inside the Black Box : The Credit Channel of Monetary Policy Transmission," *Journal of Economic Perspectives*, 9(4), 27–48.

BIGIO, S. (2013): "Endogenous Liquidity and the Business Cycle," *New York University*.

BRUNNERMEIER, B. M. K., AND Y. SANNIKOV (2014): "A Macroeconomic Model with a Financial Sector," *American Economic Review*, 104(2), 379–421.

BRUNNERMEIER, M. K. (2009): "Deciphering the Liquidity and Credit Crunch 2007-2008," *Journal of Economic Perspectives*, 23(1), 77–100.

BRUNNERMEIER, M. K., T. M. EISENBACH, AND Y. SANNIKOV (2012): "Macroeconomics with Financial Frictions: A Survey," *NBER Working Paper*, 18102.

CARLSTROM, C. T., AND T. S. FUERST (1997): "Agency Costs , Net Worth , and Business Fluctuations : A Computable General Equilibrium Analysis," *American Economic Review*, 87(5), 893–910.

CHRISTIANO, L. J., R. MOTTO, AND M. ROSTAGNO (2014): "Risk Shocks," *American Economic Review*, 104(1), 27–65.

CHRISTIANO, L. J., M. TRABANDT, AND K. WALENTIN (2011): "Introducing Financial Frictions and Unemployment into a Small Open Economy Model," *Journal of Economic Dynamics and Control*, 35(December), 1999–2041.

CLARIDA, R. H., J. GALÍ, AND M. GERTLER (2000): "Monetary Policy Rules and Macroeconomic Stability : Evidence and Some Theory," *Quarterly Journal of Economics*, 115(1), 147–180.

Covas, F., and S. Fujita (2010): "Procyclicality of Capital Requirements in a General Equilibrium Model of Liquidity Dependence," *International Journal of Central Banking*, 6(4), 137–173.

Del Negro, M., G. Eggertsson, A. Ferrero, and N. Kiyotaki (2011): "The Great Escape? A Quantitative Evaluation of the Fed's Liquidity Facilities," *Federal Reserve Bank of New York Staff Report*, 520.

Duffie, D. (2010): "The Failure Mechanics of Dealer Banks," *Journal of Economic Perspectives*, 24(1), 51–72.

Galí, J., and M. Gertler (1999): "Inflation dynamics : A structural econometric analysis," *Journal of Monetary Economics*, 44, 195–222.

Gertler, M., and P. Karadi (2011): "A model of unconventional monetary policy," *Journal of Monetary Economics*, 58(1), 17–34.

Gertler, M., and N. Kiyotaki (2011): "Financial Intermediation and Credit Policy in Business Cycle Analysis," in *Handbook of Monetary Economics*, vol. 3A, chap. 11, pp. 547–599. Elsevier Ltd.

Gorton, G., and A. Metrick (2010): "Haircuts," *Federal Reserve Bank of St. Louis Review*, 92(6), 507–519.

——— (2012): "Securitized banking and the run on repo," *Journal of Financial Economics*, 104(3), 425–451.

Gorton, G., and G. Ordoñez (2014): "Collateral Crises," *American Economic Review*, 104(2), 343–78.

Heider, F., M. Hoerova, and C. Holthausen (2009): "Liquidity Hoarding and Interbank Market Spreads: The Role of Counterparty Risk," *ECB Working Paper*, 1126.

Holmström, B., and J. Tirole (1997): "Financial Intermediation, Loanable Funds, and the Real Sector," *Quarterly Journal of Economics*, 112(3), 663–691.

——— (1998): "Private and Public Supply of Liquidity," *Journal of Political Economy*, 106(1), 1–40.

——— (2011): *Inside and Outside Liquidity.* MIT Press, Cambridge, Mass.

Kato, R. (2006): "Liquidity, infinite horizons and macroeconomic fluctuations," *European Economic Review*, 50(5), 1105–1130.

King, R. G., and S. T. Rebelo (1999): "Resuscitating Real Business Cycles," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1.

Kiyotaki, N., and J. Moore (1997): "Credit Cycles," *Journal of Political Economy*, 105(2), 211–248.

——— (2012): "Liquidity, Business Cycles, and Monetary Policy," *NBER Working Paper*, 17934.

Kurlat, P. (2013): "Lemons Markets and the Transmission of Aggregate Shocks," *American Economic Review*, 103(4), 1463–1489.

Meh, C. A., and K. Moran (2010): "The role of bank capital in the propagation of shocks," *Journal of Economic Dynamics and Control*, 34(3), 555–576.

Rotemberg, J. J. (1982): "Monopolistic Price and Adjustment Output Aggregate," *Review of Economic Studies*, 49(4), 517–531.

Tirole, J. (2011): "Illiquidity and All Its Friends," *Journal of Economic Literature*, 49(2), 287–325.

Townsend, R. M. (1979): "Optimal Contracts and Competitive with Costly State Verification," *Journal of Economic Theory*, 293, 265–293.

A Technical Appendix

A.1 Proof of Proposition 1

(i) Note that

$$\frac{\partial H(\bar{\omega})}{\partial \bar{\omega}} = f(\bar{\omega}_0)\left\{\mu + q\left[\bar{\omega}_0\left(F(\bar{\omega}_0)\pi_H R^h + (1-F(\bar{\omega}_0))\xi^*\right) - \left(1-\int_0^{\bar{\omega}_0}\omega\,\mathrm{d}F(\omega)\right)\left(\pi_H R^h - \xi^*\right)\right]\right\}$$
$$= 0$$
$$\iff \bar{\omega}_0 = \frac{\left(1-\int_0^{\bar{\omega}_0}\omega\,\mathrm{d}F(\omega)\right)\left(\pi_H R^h - \xi^*\right) - \frac{\mu}{q}}{F(\bar{\omega}_0)\pi_H R^h + (1-F(\bar{\omega}_0))\xi^*}$$

Then, we have that

$$\frac{\partial H(\bar{\omega})}{\partial \bar{\omega}} = \begin{cases} < 0 : \bar{\omega} < \bar{\omega}_0 \\ = 0 : \bar{\omega} = \bar{\omega}_0 \\ > 0 : \bar{\omega} > \bar{\omega}_0 \end{cases}$$

i.e. $H(\bar{\omega})$ achieves a minimum at $\bar{\omega}_0$. Since the loan scale (5) is a function of $H(\bar{\omega})^{-1}$, it is maximised at this point. Furthermore, F increases monotonically in $\bar{\omega}$. Hence, bankers can always improve on $\bar{\omega} < \bar{\omega}_0$ by choosing $\bar{\omega} = \bar{\omega}_0$. The latter must, therefore, be the lower bound of the liquidity threshold.

To relate $\bar{\omega}_0$ to the interpretation as the indifference threshold mentioned in section 2.3.1, note that

$$q\pi_H R^h l - \left(1-\int_0^{\bar{\omega}_0}\omega\,\mathrm{d}F(\omega)\right)^{-1}(\bar{\omega}_0 d + \mu l) = q\xi^* l$$
$$\iff \bar{\omega}_0 = \left[\left(1-\int_0^{\bar{\omega}_0}\omega\,\mathrm{d}F(\omega)\right)\left(\pi_H R^h - \xi^*\right)q - \mu\right]\frac{l}{d}$$
$$\iff \bar{\omega}_0 = \frac{\left(1-\int_0^{\bar{\omega}_0}\omega\,\mathrm{d}F(\omega)\right)\left(\pi_H R^h - \xi^*\right) - \frac{\mu}{q}}{F(\bar{\omega}_0)\pi_H R^h + (1-F(\bar{\omega}_0))\xi^*}$$

where the last step uses the participation constraint of households, (2).

(ii) Completing the derivation of the first-best refinancing threshold in section 2.3 by again using households' participation constraint (2) yields

$$\begin{aligned} q\pi_H Rl - \bar{\omega}_1 d &= q\xi^* l \\ \Longleftrightarrow \bar{\omega}_1 &= (\pi_H R - \xi^*) \frac{ql}{d} \\ \Longleftrightarrow \bar{\omega}_1 &= \frac{\pi_H R - \xi^*}{F(\bar{\omega}_1)\pi_H R + (1 - F(\bar{\omega}_1))\xi^*} \end{aligned}$$

(iii) Finally, (6) is convex in $\bar{\omega}$, such that the necessary and sufficient condition for a maximum is given by the first order condition (9).[29]

□

A.2 Properties of the Optimal Liquidity Threshold

First determine the partial derivatives of Q with respect to $\bar{\omega}$, q, and ξ^*:

$$\frac{\partial Q}{\partial \bar{\omega}} = q \left\{ \bar{\omega} f(\bar{\omega}) \left[2F(\bar{\omega}) \left(\pi_H R^h - \xi^* \right) \right] + F(\bar{\omega}) \left(F(\bar{\omega}) \pi_H R^h + (1 - F(\bar{\omega}))\, \xi^* \right) \right\}$$

A sufficient condition for $\frac{\partial Q}{\partial \bar{\omega}} > 0$ is $\pi_H R^h > \xi^*$, which will be satisfied in the calibration. Next,

$$\frac{\partial Q}{\partial \xi^*} = 1 - \int_0^{\bar{\omega}} \omega \, dF(\omega) + \bar{\omega} F(\bar{\omega}) \left(1 - F(\bar{\omega}) \right) > 0$$

since the share of deposits that are not buffered by liquid reserves $\int_0^{\bar{\omega}} \omega \, dF(\omega) \leq 1$ by definition. Finally

$$\frac{\partial Q}{\partial q} = \left(1 - \int_0^{\bar{\omega}} \omega \, dF(\omega) \right) \xi^* + \bar{\omega} F(\bar{\omega}) \left(F(\bar{\omega}) \pi_H R^h + (1 - F(\bar{\omega}))\, \xi^* \right) > 0$$

for the same reason. Now applying the implicit function theorem (IFT) we have that

$$\frac{\partial \bar{\omega}}{\partial \xi} = - \frac{\partial Q / \partial \xi}{\partial Q / \partial \bar{\omega}} < 0$$

and

$$\frac{\partial \bar{\omega}}{\partial q} = - \frac{\partial Q / \partial q}{\partial Q / \partial \bar{\omega}} < 0$$

.

[29] In the numerical exercise I will show that, indeed, $\bar{\omega}_0 < \bar{\omega} < \bar{\omega}_1$.

A.3 Dynamic Equilibrium Conditions

Given the aggregate state variables $\Gamma = \left(K_t, K_t^b, B_t; z_t, z_t^{\xi}\right)$, the competitive equilibrium is a set of policy functions pinning down

$$\left(K_{t+1}, K_{t+1}^b, K_{t+1}^h, B_{t+1}, C_t^b, C_t^h, I_t, L_t, A_t, H_t^b, H_t^h, R_t, R_t^b, R_t^h, q_t, r_t, r_t^b, w_t^b, w_t^h\right)$$

together with the exogenous laws of motion of $\left(z_t, z_t^{\xi}\right)$. The solution to the dynamic programming problem satisfies the following set of equilibrium conditions

1. Individual optimality

   (a) Households

$$1 = \beta E_t \left[\left(\frac{C_{t+1}^h}{C_t^h} \right)^{-\theta} \frac{q_{t+1}(1-\delta) + r_{t+1}}{q_t} \right] \tag{A.1}$$

$$1 = \beta E_t \left[\left(\frac{C_{t+1}^h}{C_t^h} \right)^{-\theta} (1 + r_{t+1}^b) \right] \tag{A.2}$$

$$-\frac{\nu}{\eta^h - H_t^h} = -C_t^{h^{-\theta}} w_t^h \tag{A.3}$$

   (b) Final good producers

$$r_t = \exp(z_t) \alpha^k K_t^{\alpha^k - 1} H_t^{h^{\alpha^h}} H_t^{b^{\alpha^b}} \tag{A.4}$$

$$w_t^b = \exp(z_t) \alpha^b K_t^{\alpha^k} H_t^{h^{\alpha^h}} H_t^{b^{\alpha^b - 1}} \tag{A.5}$$

$$w_t^h = \exp(z_t) \alpha^h K_t^{\alpha^k} H_t^{h^{\alpha^h - 1}} H_t^{b^{\alpha^b}} \tag{A.6}$$

   (c) Banks

$$R_t^b = \frac{\mu}{q_t(\pi_H - \pi_L)} \tag{A.7}$$

$$R_t = R_t^b + R_t^h \tag{A.8}$$

$$L_t = \frac{A_t}{H(\bar{\omega}_t)} \tag{A.9}$$

$$1 = Q(\bar{\omega}_t) \tag{A.10}$$

$$A_t = (q_t(1-\delta) + r_t) K_t^b + H_t^b w_t^b \tag{A.11}$$

$$K_{t+1}^b = \tau^b F(\bar{\omega}_t) \pi_H R_t^b L_t \tag{A.12}$$

$$C_t^b = (1 - \tau^b) q_t F(\bar{\omega}_t) \pi_H R_t^b L_t \tag{A.13}$$

$$H_t^b = \eta^b \tag{A.14}$$

where

$$H(\bar{\omega}_t) = 1 + F(\bar{\omega}_t)\mu - q_t \left(1 - \int_0^{\bar{\omega}_t} \omega \, dF(\omega)\right) \left(F(\bar{\omega}_t)\pi_H R^h + (1 - F(\bar{\omega}_t))\, \xi_t^*\right)$$
$$Q(\bar{\omega}_t) = q_t \left[\left(1 - \int_0^{\bar{\omega}_t} \omega \, dF(\omega)\right) \xi_t^* + \bar{\omega}_t F(\bar{\omega}_t) \left(F(\bar{\omega}_t)\pi_H R^h + (1 - F(\bar{\omega}_t))\, \xi_t^*\right)\right]$$
$$\xi_t^* = \xi + z_t^\xi$$

2. Market clearing conditions

(a) Goods

$$C_t^h + q_t K_{t+1}^h + B_{t+1} = (1 + r_t^b) B_t + (q_t(1-\delta) + r_t) K_t^h + w_t^h H_t^h \tag{A.15}$$

(b) Capital

$$K_t = K_t^b + K_t^h \tag{A.16}$$

(c) Investment

$$K_{t+1} = (1-\delta)K_t + (F(\bar{\omega}_t)\pi_H R + (1 - F(\bar{\omega}_t))\xi)\, I_t \tag{A.17}$$

(d) Loans

$$q_t L_t = q_t I_t \tag{A.18}$$

(e) Bonds

$$B_t = 0 \tag{A.19}$$

3. Exogenous processes

$$z_t = \rho z_{t-1} + e_t \tag{A.20}$$
$$z_t^\xi = \rho_\xi z_{t-1}^\xi + e_t^\xi \tag{A.21}$$

A.4 Steady State

Endogenous parameters $\{\nu, R, \mu, \sigma^2(\omega), \tau^b, \xi\}$ are solved from the following calibration targets:

- ν: $H^h = 0.3$
- R: $\pi_H R = 1$

- $\{\mu, \sigma^2(\omega), \tau^b, \xi\}$:

$$
\begin{aligned}
LGD &= 1 - \frac{q\xi^* I}{I - A} \\
\Xi &= \frac{D}{A} \\
\Omega &= \int_0^{\bar{\omega}} \omega \, dF(\omega) \frac{D}{A + D} = \int_0^{\bar{\omega}} \omega \, dF(\omega) \left(1 + \Xi^{-1}\right)^{-1} \\
\Phi &= \left[F(\bar{\omega})\pi_H R^h + (1 - F(\bar{\omega}))\,\xi\right] \frac{I}{Y}
\end{aligned}
$$

where $D = q\left[F(\bar{\omega})\pi_H R^h + (1 - F(\bar{\omega}))\,\xi\right] L$ and $Y = K^{\alpha^k} H^{h^{\alpha^h}} H^{b^{\alpha^b}}$. Further, let $\omega \sim U[0, U_H]$ such that $\sigma^2(\omega) = \frac{U_H^2}{12}$. Given these targets, the steady state can be derived as follows:

$$
\begin{aligned}
B &= 0 \\
R &= \frac{1}{\pi_H} \\
r^b &= \frac{1}{\beta} - 1 \\
H^b &= \eta^b
\end{aligned}
$$

To continue, guess $\{q, \bar{\omega}, \mu, \xi\}$, then:

$$
\begin{aligned}
F(\bar{\omega}) &= 2\left(1 + \Xi^{-1}\right) \Omega \bar{\omega}^{-1} \\
r &= \left(\frac{1}{\beta} - (1 - \delta)\right) q \\
R^b &= \frac{\mu}{q(\pi_H - \pi_L)} \\
R^h &= R - R^b \\
I &= L \\
\frac{A}{L} &= 1 - \frac{q\xi}{1 - LGD} \\
D &= q\left[F(\bar{\omega})\pi_H R^h + (1 - F(\bar{\omega}))\,\xi\right] L \\
U_H &= \frac{\bar{\omega}^2}{2}\left(1 + \frac{A}{D}\right)^{-1} \Omega^{-1} \\
K &= \left(r^{-1}\alpha^k H^{h^{\alpha^h}} H^{b^{\alpha^b}}\right)^{\frac{1}{1-\alpha^k}} \\
L &= \frac{\delta}{F(\bar{\omega}_t)\pi_H R + (1 - F(\bar{\omega}_t))\xi} K \\
A &= \left(1 - \frac{q\xi}{1 - LGD}\right) L
\end{aligned}
$$

$$
\begin{aligned}
w^b &= \alpha^b K^{\alpha^k} H^{h^{\alpha^h}} H^{b^{\alpha^b - 1}} \\
w^h &= \alpha^h K^{\alpha^k} H^{h^{\alpha^h - 1}} H^{b^{\alpha^b}} \\
K^b &= (q(1-\delta) + r)^{-1} \left(A - w^b H^b\right) \\
K^h &= K - K^b \\
\tau^b &= K^b \left(F(\bar{\omega})\pi_H R^b L\right)^{-1} \\
C^b &= (1-\tau^b) q F(\bar{\omega}) \pi_H R^b L \\
C^h &= (r - q\delta) K^h + w^h H^h \\
\nu &= C^{h^{-\theta}} w^h \left(\eta^h - H^h\right)
\end{aligned}
$$

To verify the guess, check the following equations for consistency:

$$
\begin{aligned}
\Xi &= q \left[F(\bar{\omega})\pi_H R^h + (1 - F(\bar{\omega}))\,\xi\right] \frac{L}{A} \\
L &= \frac{A}{H(\bar{\omega})} \\
1 &= Q(\bar{\omega}) \\
\Phi &= \left[F(\bar{\omega})\pi_H R^h + (1 - F(\bar{\omega}))\,\xi\right] \frac{I}{Y}
\end{aligned}
$$

A.5 Derivation of the NK Phillips Curve

In the extended model, final good producers assemble intermediate goods via a Dixit-Stiglitz aggregator

$$
Y_t = \left[\int_0^1 Y_{it}^{\frac{\epsilon-1}{\epsilon}} \,\mathrm{d}i\right]^{\frac{\epsilon}{\epsilon-1}}
$$

where ϵ is the elasticity of substitution between different varieties of intermediate goods. From the optimization problem of final good producers demand for intermediate goods is given by

$$
Y_{it} = \left[\frac{P_{it}}{P_t}\right]^{-\epsilon} Y_t
$$

where P_t is the aggregate price level and P_{it} the price of variety i. Demand from intermediate producers thus depends on the relative price of their product as well as the elasticity of substitution.

Intermediate producers use capital and labour as inputs into their production function

$$
Y_{it} = \exp(z_t) F(K_{it}, H_{it}^h, H_{it}^b)
$$

Their optimization problem can be broken down into two separate steps: a cost minimization step in the production of a given quantity of intermediate goods and a price setting step. To minimize costs, intermediate producers solve

$$\min_{\{K_{it},H^h_{it},H^b_{it}\}} r_t K_{it} + w^h_t H^h_{it} + w^b_t H^b_{it} - mc_{it}[exp(z)F(K_{it},H^h_{it},H^b_{it}) - Y_{it}]$$

where the Lagrange multiplier represents the marginal cost of the firm. Taking the first order conditions and imposing symmetry across firms yields

$$r_t = mc_t \exp(z_t) F_K(K_t, H^h_t, H^b_t) \tag{A.22}$$

$$w^h_t = mc_t \exp(z_t) F_{H^h}(K_t, H^h_t, H^b_t) \tag{A.23}$$

$$w^b_t = mc_t \exp(z_t) F_{H^b}(K_t, H^h_t, H^b_t) \tag{A.24}$$

In a second step, intermediate producers set their optimal relative price given quadratic price adjustment costs subject to their individual demand schedule

$$\max_{\{P_{it}\}} \mathbb{E}_t \sum_{s=t}^{\infty} \Lambda_{s,t} \left\{ \left[\frac{P_{is}}{P_s} - mc_s\right] Y_{is} - \frac{\chi}{2}\left[\frac{P_{is}}{\bar{\pi}P_{is-1}} - 1\right]^2 Y_s \right\}$$

subject to $Y_{is} = \left[\frac{P_{is}}{P_s}\right]^{-\epsilon} Y_s$ and where $\Lambda_{s,t} = \frac{\beta\lambda_s}{\lambda_t}$ is households' stochastic discount factor between periods s and t. In a symmetric equilibrium all intermediate good producers set the same price, such that type subscript i can be dropped and $P_{it} = P_t$, $\forall i \in [0,1]$. After some manipulations, the first order condition yields the forward-looking New Keynesian Phillips curve

$$\left(\frac{\pi_t}{\bar{\pi}} - 1\right)\left(\frac{\pi_t}{\bar{\pi}}\right) = \frac{\epsilon}{\chi}\left(m_{c,t} - \frac{\epsilon-1}{\epsilon}\right) + \mathbb{E}_t\left[\frac{\beta\lambda_{t+1}}{\lambda_t}\left(\frac{\pi_{t+1}}{\bar{\pi}} - 1\right)\left(\frac{\pi_{t+1}}{\bar{\pi}}\right)\frac{Y_{t+1}}{Y_t}\right] \tag{A.25}$$

The second sector directly affected by the introduction of nominal rigidities is the household sector. Households now choose the level of nominal rather than real bonds

$$\max_{\{c^h_t,k^h_{t+1},b^h_{t+1},h^h_t\}} E_0 \sum_{t=0}^{\infty} \beta^t u(c^h_t, h^h_t)$$

$$s.t.$$

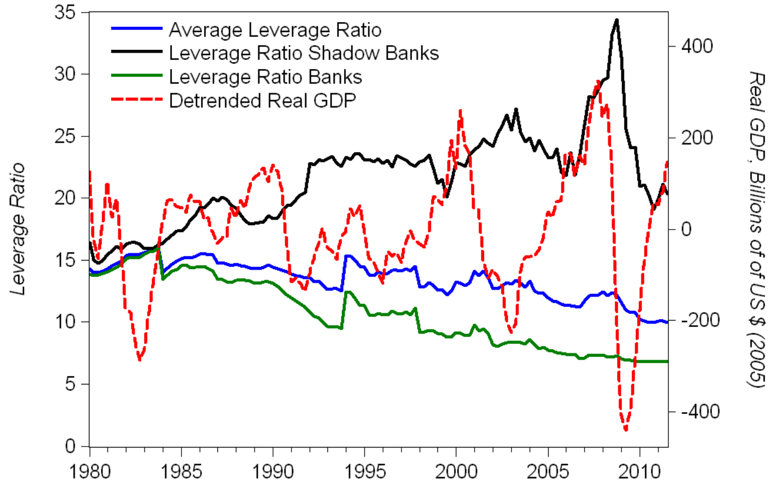
$$c^h_t + q_t k^h_{t+1} + \frac{b^h_{t+1}}{P_t} = (1+i^b_t)\frac{b^h_t}{P_t} + (q_t(1-\delta) + r_t)k^h_t + w^h_t h^h_t \tag{A.26}$$

and the first order condition for b^h_{t+1} accordingly becomes

$$\lambda_t = \beta E_t\left[\lambda_{t+1}\frac{(1+i^b_{t+1})}{\pi_{t+1}}\right] \tag{A.27}$$

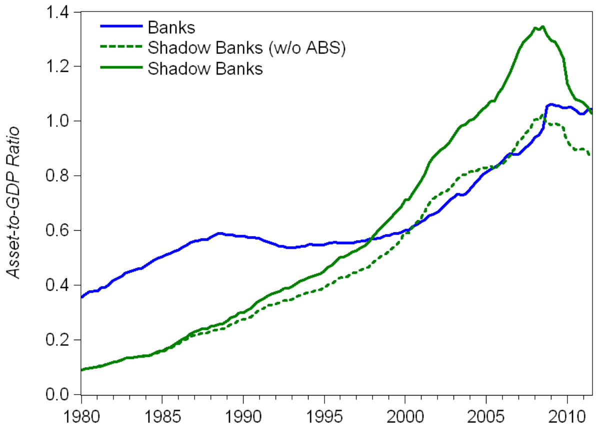
B Graphs

Figure 6: Leverage Ratios of Banks and Market-based Intermediaries



Notes: US-chartered commercial banks, savings institutions and credit unions are identified as traditional banks. The shadow banking sector comprises securities and broker dealers, issuers of asset-backed securities, finance companies and Government-sponsored enterprises. This follows the classification in Adrian and Shin (2009). The leverage ratio is defined as the ratio of debt to equity. Source: US Flow of Funds (Federal Reserve)

Figure 7: Asset-to-GDP Ratios of Banks and Market-based Intermediaries



Source: US Flow of Funds (Federal Reserve)