Wholesale Banking and Bank Runs in Macroeconomic Modelling of Financial Crises

Mark Gertler, Nobuhiro Kiyotaki and Andrea Prestipino*
NYU, Princeton University and Federal Reserve Board of Governors

April 2015

*Prepared for Handbook of Macroeconomics, edited by John B. Taylor and Harald Uhlig.

1 Introduction

One of the central challenges for contemporary macroeconomics is adapting the core models to account for why the recent financial crisis occurred and why it then devolved into the worst recession of the postwar period. On the eve of the crisis the basic workhorse quantitative models used in practice largely abstracted from financial market frictions. These models were thus largely silent on how the crisis broke out and how the vast array of unconventional policy interventions undertaken by the Federal Reserve and Treasury could have worked to mitigate the effects of the financial turmoil. Similarly, these models could not provide guidance for the regulatory adjustments needed to avoid another calamity.

From the start of the crisis there has been an explosion of literature aimed at meeting this challenge. Much of the early wave of this literature builds on the financial accelerator and credit cycle framework developed in Bernanke and Gertler (1989) and Kiyotaki and Moore (1997). This approach stresses the role of balance sheets in constraining borrower spending in a setting with financial market frictions. Procyclical movement in balance sheet strength amplifies spending and thus aggregate economic activity. A feedback loop emerges as conditions in the real economy affect the condition of balance sheets and vice-versa. Critical to this mechanism is the role of leverage: The exposure of balance sheets to systemic risk is increasing in the degree of borrower leverage.

The new vintage of macroeconomic models with financial frictions makes progress in two directions: First, it adapts the framework to account for the distinctive features of the current crisis. In particular, during the recent crisis, it was highly leveraged financial institutions along with highly leveraged households that were most immediately vulnerable to financial distress. The conventional literature featured balance sheet constraints on non-financial firms. Accordingly, a number of recent macroeconomic models have introduced balance sheet constraints on banks, while others have done so for households.[1] The financial accelerator remains operative, but the class of agents most directly affected by the financial market disruption differ from earlier work.

Another direction has involved improving the way financial crises are

[1] See Gertler and Karadi (2012), Gertler and Kiyotaki (2010) and Curdia and Woodford (2010) for papers that incorporate banking and Eggertsson and Krugman (2012) and Geurreri and Lorenzoni (2011) for papers that incuded household debt.

modeled. For example, financial crises are inherently nonlinear events, often featuring a simultaneous sudden collapse in asset prices and rise in credit spreads.[2] A sharp collapse in output typically ensues. Then recovery occurs only slowly, as it is impeded by a slow process of deleveraging. A number of papers have captured this nonlinearity by allowing for the possibility that the balance sheet constraints do not always bind.[3] Financial crises are then periods where the constraints bind, causing an abrupt contraction in economic activity. Another approach to handling the nonlinearity is to allow for bank runs.[4] Indeed, runs on the shadow banking system were a salient feature of the crisis, culminating with the collapse in September 2008 of Lehman Brothers, of some major money market funds and ultimately of the entire investment banking sector. Yet another literature capture the nonlinearity inherent in financial crises by modeling network interactions.

One area the macroeconomics literature has yet to address adequately is the distinctive role of the wholesale banking sector in the breakdown of the financial system. Our notion of wholesale banks corresponds roughly, though not exactly, to the shadow banking sector on the eve of the 2007-2009 financial crisis. Shadow banking includes all financial intermediaries that operated outside the Federal Reserve's regulatory framework. By wholesale banking, we mean the subset that (i) was highly leveraged, often with short term debt and (ii) relied heavily on borrowing from other financial institutions in "wholesale" markets for bank credit, as opposed to borrowing from households in "retail" markets for bank credit.

When the crisis hit, the epicenter featured malfunctioning of the wholesale banking sector. Indeed, retail markets remained relatively stable while wholesale funding markets experienced dry-ups and runs. By contrast, much of the recent macroeconomic modeling of banking features traditional retail banking. In this respect it misses some important dimensions of both the run-up to the crisis and how exactly the crisis played out. In addition, by omitting wholesale banking, the literature may be missing some important considerations for regulatory design.

In this Handbook chapter we present a simple canonical macroeconomic

[2]See Krishnamurthy, Nagel and Orlov (2014) for evidence in support of the nonlinearity of financial crises.

[3]See Brunnermeier and Sannikov (2014), He and Krishnamurthy (2013,2014) and Mendoza (2010).

[4]See Gertler and Kiyotaki (2015 forthcoming), Ferrante (2015), Robatto (2014) and Martin, Skeie and Von Thadden (2014a,b).

model of banking crises that (i) is representative of the existing literature; and (ii) extends this literature to feature a role for wholesale banking. The model will provide some insight both into the growth of wholesale banking and into how this growth led to a build-up of financial vulnerabilities that ultimately led to a collapse. Because the model builds on existing literature, our exposition of the framework will permit us to review the progress that is made. However, by turning attention to wholesale banks and wholesale funding markets, we are able to chart a direction we believe the literature should take.

In particular, the model is an extension of the framework developed in Gertler and Kiyotaki (2011), which had a similar two-fold objective: first, present a canonical framework to review progress that has been made and, second, chart a new direction. That paper characterized how existing financial accelerator models that featured firm level balance sheet constraints could be extended to banking relationships in order to capture the disruption of banking during the crisis. The model developed there considered only retail banks which funded loans mainly from household deposits. While it allowed for an inter-bank market for credit among retail banks, it did not feature banks that relied primarily on wholesale funding, as was the case with shadow banks.

For this Handbook chapter we modify the Gertler and Kiyotaki framework to incorporate wholesale banking alongside retail banking, where the amount credit intermediated via wholesale funding markets arises endogenously. Another important difference is that we allow for the possibility of runs on wholesale banks. We argue that both these modifications improve the ability of macroeconomic models to capture how the crisis played out. They also provide insight into how the financial vulnerabilities built up in the first place.

As way to motivate our emphasis on wholesale banking, Section 2 present some descriptive evidence on both the growth of this sector and the collapse it experienced during the Great Recession. It also describes how the collapse contributed to the downturn. Section 3 presents the baseline macroeconomic model with banking, where a wholesale banking sector arises endogenously. We also illustrate how runs that significantly disrupt the economy are possible in wholesale funding markets. Sector 4 conducts a set of numerical experiments. First we show how technological improves that work to reduce agency frictions in wholesale funding markets can account for the growth of a highly leveraged wholesale banking sector. While the increased size of the wholesale banking improves the efficiency of financial intermediation, it also raises the vulnerability of this sector to runs. We illustrate this point with some numerical experiments. Section 5 considers the case where runs in the wholesale sector might be anticipated. Here we illustrate how the anticipation of a run can have harmful effects. It induces an increase in credit spreads and a kind of "slow" run, where creditors draw down deposits, similar to what occurred in the recent crisis prior to the collapse of Lehman Brothers. The section also illustrates how the model can capture some of the key phases of the financial collapse, including both the slow run period up to Lehman and the ultimate "fast run" collapse. In Section 6 we discuss policy. Section 7 review the literature. Concluding remarks are in Section 8.

2 The Growth and Fragility of Wholesale Banking

In this section we provide some background motivation for the canonical macroeconomic model with wholesale bank funding markets that we develop in the following section. We do so by presenting a brief description of the growth and ultimate collapse of wholesale funding markets during the Great Recession. We also describe informally how the disruption of these markets contributed to the contraction of the real economy.

Figure 1 illustrates how we consider the different roles of retail and wholesale financial intermediaries, following the tradition of Gurley and Shaw (1960).[5] The arrows indicate the direction that credit is flowing. Funds can flow from households (ultimate lenders) to non-financial borrowers (ultimate borrowers) through three different paths: they can be lent directly from households to borrowers $\left(K^h\right)$; they can be intermediated by retail

[5] Gurley and Show (1960) consider that there are two ways to transfer fund from ultimate lenders (with surplus funds) to ultimate borrowers (who need external funds to finance expenditure): direct and indirect finance. In direct finance, ultimate borrowers sell their securities directly to ultimate lenders to raise fund. In indirect finance, financial intermediaries sell their own securities to raise fund from ultimate lenders in order to buy securities from ultimate borrowers. By doing so, financial intermediaries transform relatively risky, illiquid and long maturity securities of ultimate borrowers into relatively safe, liquid and short maturity securities of intermediaries. Here we divide financial intermediaries into wholesale and retail financial intermediaries, while both involve asset transformation of risk, liquidity and maturity. We call intermediaries "banks" and ultimate lenders as "households" for short.

banks that raise deposits (D) from households and use them to make loans to non-financial borrowers (K^r); alternatively, lenders' deposits can be further intermediated by specialized financial institutions that raise funds from retail banks in wholesale funding markets (B) and, in turn, make loans to ultimate borrowers (K^w). In what follows we refer to these specialized financial institutions as wholesale banks.

We think of wholesale banks as highly leveraged shadow banks that rely heavily on credit from other financial institutions, particularly short term credit. We place in this category institutions that financed long term assets such as mortgaged back securities with short term money market instruments, including commercial paper and repurchase agreements. Examples of these kinds of financial institutions are investment banks, hedge funds and conduits. We focus attention on institutions that relied heavily on short term funding in wholesale markets to finance longer term assets because it was primarily these kinds of entities that experienced financial turmoil.

Our retail banking sector, in turn, includes financial institutions that rely mainly on household saving for external funding and provide a significant amount of short term financing to the wholesale banks. Here we have in mind commercial banks, money market funds and mutual funds that raised funds mainly from households and on net provided financing to wholesale banks.

The conventional macroeconomic models of banking ignore the flow of intermediation via wholesale banks and instead focus only on credit intermediated by retail banks. By doing so, however, they miss the chain through which the crisis was propagated. As we discuss shortly, the crisis began with defaults on mortgages held by wholesale banks which ultimately led to a collapse in markets for wholesale funding. Eventually the turmoil was felt in the traditional retail banking sector as well, but the origin of the crisis was in wholesale banking.

Figure 1 treats wholesale banking as if it is homogenous. In order to understand how the crisis spread, it is useful to point out that there are different layers within the wholesale banking sector. While the intermediation process was rather complex, conceptually we can reduce the number of layers to three basic ones: (1) origination; (2) securitization; (3) and funding. Figure 2 illustrates the chain. First there are "loan originators," such as mortgage origination companies and finance companies, that made loans directly to non-financial borrowers. At the other end of the chain were shadow banks that held securitized pools of the loans made by originators (e.g. asset backed security issuers). In between were brokers and conduits that assisted in the securitization process and provided market liquidity. Dominant in this group were the major investment banks (e.g., Goldman Sachs, Morgan Stanley, Lehman Brothers, etc.). Each of these layers relied on short term funding, including commercial paper, asset-backed commercial paper and repurchase agreements. While there was considerable inter-bank lending among wholesale banks, retail banks (particularly money market funds) on net provided short term credit in wholesale credit markets.

We next describe a set of facts about wholesale banking. We emphasize three sets of facts in particular: (1) wholesale banking grew in relative importance over the last four decades; (2) leading up to the crisis wholesale banks were highly exposed to systemic risk because they were highly leveraged and relied heavily on short term debt; and (3) the subsequent disruption of wholesale funding markets raised credit costs and contracted credit flows, likely contributing in a major way to the Great Recession.

1. *Growth in Wholesale Banking*

We now present measures of the scale of wholesale banking relative to retail banking as well as to household's direct asset holdings. Table 1 describes how we construct measures of assets held by wholesale versus retail banks. In particular it lists how we categorized the various types of financial intermediaries into wholesale versus retail banking.[6,7] As the table indicates, the wholesale banking sector aggregates financial institutions that originate loans, that help securitize them and that ultimately fund them. A common feature of all these institutions, though, is that they relied heavily on short

[6]The Appendix provides details about measurement of the time series shown in this section from Flow of Funds data.

[7]What is important to notice is that the measures we report are robust to alternative approaches. See, e.g., Adrian and Ashcraft 2012 NYFed Report, for an alternative definition of shadow banking that yields very simlar conclusions and Pozsar et al 2013 NYFed Report, for a detailed description of shadow banking.

term credit in wholesale funding markets.

Table 1		
Retail Sector	Private Depository Institutions Money Market Mutual Funds Mutual Funds	
Wholesale Sector	*Origination*	Finance Companies Real Estate Investment Trusts Government Sponsored Enterprises
	Securitization	Security Brokers Dealers ABS Issuers
	Funding	GSE Mortgage Pools Funding Corporations Holding Companies

Figure 3 portrays the log level of assets held by wholesale banks, by retail banks, and directly by households from the early 1980s until the present. The figure shows the rapid increase in wholesale banking relative to the other means of asset holding. Wholesale banks went from holding under fifteen percent of total assets in the early 1980s to roughly forty percent on the eve of the Great Recession, an amount on par with assets held by retail banks.

Two factors were likely key to the growth of wholesale banking. The first is regulatory arbitrage. Increased capital requirements on commercial banks raised the incentive to transfer asset holding outside the commercial bank system. Second, financial innovation improved the liquidity of wholesale funding markets. The securitization process in particular improved (perceived) safety of loans by diversifying idiosyncratic risks as well as by enhancing the liquidity of secondary markets for bank assets. The net effect was to raise the borrowing capacity of the overall financial intermediary sector.

2. *Growth in Leverage and Short Term Debt in Wholesale Banking*

Wholesale banking not only grew rapidly, it also became increasingly vulnerable to systemic disturbances. Figure 4 presents evidence on the growth in leverage in the investment banking sector. Specifically it plots the aggregate leverage multiple for broker dealers (primarily investment banks) from 1980 to the present. We define the leverage multiple as the ratio of total assets held to equity.[8] The greater is the leverage multiple, the higher is the

[8] The data is from Flow of Funds and equity is measured by book value.

reliance on debt finance relative to equity. The key takeaway from Figure 4 is that the leverage multiple grew from under five in the early 1980s to over forty at the beginning of the Great Recession, a nearly tenfold increase.

Arguably, the way securitization contributed to the overall growth of wholesale banking was by facilitating the use of leverage. By constructing assets that appeared safe and liquid, securitization permitted wholesale banks to fund these assets by issuing debt. At a minimum debt finance had the advantage of being cheaper due to the tax treatment. Debt financing was also cheaper to the extent the liabilities were liquid and thus offered a lower rate due to a liquidity premium.

Why were these assets funded in wholesale markets as opposed to retail markets? The sophistication of these assets required that creditors be highly informed to evaluate payoffs, especially given the absence of deposit insurance. The complicated asset payoff structure also suggests that having a close working relationship with borrowers is advantageous. It served to reduce the possibility of any kind of financial malfeasance. Given these considerations, it makes sense that wholesale banks obtain funding in inter-bank markets. In these markets lenders are sophisticated financial institutions as opposed to relatively unsophisticated households in the retail market.

Figure 5 shows that much of the growth in leverage in wholesale banking involved short term borrowing. The figure plots the log levels of asset backed commercial paper (ABCP) and repurchase agreements (Repo). This growth reflected partly the growth in assets held by wholesale banks and partly innovation in loan securitization that made maturity transformation by wholesale banks more efficient. Also relevant, however, was a shift in retail investors demand from longer term security tranches towards short term credit instruments as the initial fall in housing prices in 2006 raised concerns about the quality of existing securitized assets.[9,10] As we discuss next, the combination of high leverage and short term debt is what made the wholesale banking system extremely fragile.

3. The Crisis: The Unraveling of Wholesale Bank Funding Markets

[9] See Brunnermeier and Oemke (2013) for a model in which investors prefer shorter maturities when realease of information could lead them not to roll over debt.

[10] It is not easy to gather direct evidence on this from the aggregate composition of liabilities of wholesale banks since data from the Flow of Funds excludes the balance sheets of SIVs and CDOs from the ABS Issuers category. Our narrative is based on indirect evidence coming from ABX spreads as documented for example in Gorton 2009 AER PP

The losses suffered by mortgage originators due to falling housing prices in 2006 eventually created strains in wholesale funding markets. Short term wholesale funding markets started experiencing severe turbulence in the summer of 2007. In July 2007 two Bear Sterns hedge funds that had invested in subprime related products declared bankruptcy. Shortly after, BNP Paribas had to suspend withdrawals from investment funds with similar exposure. These two episodes led investors to reassess the risks associated with the collateral backing commercial paper offered by asset backed securities issuers. In August 2007 a steady contraction of Asset Backed Commercial Paper (ABCP) market began, something akin to a "slow run", in Bernanke's terminology.[11] The value of Asset Backed Commercial Paper outstanding went from a peak of 1.2 trillion dollars in July 2007 to 800 billion dollars in December of the same year and continued its descent to its current level of around 200 billion dollars.

The second significant wave of distress to hit wholesale funding markets featured the collapse of Lehman Brothers in September of 2008. Losses on short term debt instruments issued by Lehman Brothers led the Reserve Primary Fund, a large Money Market Mutual Fund (MMMF), to "break the buck": the market value of assets fell below the value of its non-contingent liabilities. An incipient run on MMMFs was averted only by the extension of Deposit Insurance to these types of institutions. Wholesale investors,[12] however, reacted by pulling out of the Repo market, switching off the main source of funding for Security Broker Dealers. Figure 5 shows the sharp collapse in repo financing around the time of the Lehman collapse. Indeed if the first wave of distress hitting the ABCP market had the features of a "slow run", the second, which led to the dissolution of the entire investment banking system had the features of a traditional "fast run."

We emphasize that a distinctive feature of these two significant waves of financial distress is that they did not involve traditional banking institutions. In fact, the retail sector as a whole was shielded thanks to prompt government intervention that halted the run on MMMFs in 2008 as well as the Troubled Asset Relief Program and other subsequent measures that supplemented the

[11] Covitz, Liang and Suarez (2013) provide a detailed description of the run on ABCP programs in 2007. A very clear description of the role of commercial paper during the 2007-2009 crisis is presented by Kacperczyk and Schnabl (2010).

[12] The poor quality of available data makes it difficult to exactly identify the identity of the investors running on Repo's. See Gorton (2010) and Krishnamurthy Nagel and Orlov (2014).

traditional safety net. In fact, total short term liabilities of the retail sector were little affected overall (See Figure 6). This allowed the retail banking sector to help absorb some of the intermediation previously performed by wholesale banks.

Despite the unprecedented nature and size of government intervention and the partial replacement of wholesale intermediation by retail bank lending, the distress in wholesale bank funding markets led to widespread deterioration in credit conditions. Figure 7 plots behavior of three representative credit spreads: (1) The spread between the three month ABCP rate and three month Treasury spread; (2) The financial company commercial paper spread; and (3) The Gilchrist and Zakrajsek (2012) excess bond premium. In each case the spread is the difference between the respective rate on the private security and a similar maturity treasury security rate. The behavior of the spreads lines up with the waves of financial distress that we described. The ABCP spread jumps 1.5% in August 2007, the beginning of the unraveling of this market. The increase in this spread implies a direct increase in credit costs for borrowing funded by ABCP including mortgages, car loans, and credit card borrowing. As problems spread to broker dealers, the financial commercial paper spread increases reaching a peak at more than 1.5% at the time of the Lehman. Increasing costs of credit for these intermediaries, in turn, helped fuel increasing borrowing costs for non-financial borrowers. The Gilchrist and Zakrajsek's corporate excess bond spread jumps more than 2.5% from early 2007 to the peak in late 2008.

It is reasonable to infer that the borrowing costs implied by the increased credit spreads contributed in an important way to the slowing of the economy at the onset of the recession in 2007:Q4, as well as to the sharp collapse following the Lehman failure. Figure 8 shows the evolution of quantity index of business investment, residential investment, durable consumption and their sum - total investment, starting from the first quarter of 2008.

In our view, there are three main conclusions to be drawn from the empirical evidence presented in this section. First, the wholesale banking sector grew into a very important component of financial intermediation by relying on securitization to reduce the risks of lending and expand the overall borrowing capacity of the financial system. Second, higher borrowing capacity came at the cost of increased fragility as high leverage made wholesale banks' net worth very sensitive to corrections in asset prices. Third, the disruptions in wholesale funding markets that took place in 2007 and 2008 seem to have played an important role in the unfolding of the Great Recession. These observations motivate our modeling approach below and our focus on interbank funding markets functioning and regulation.

3 Basic Model

3.1 Key Features

Our starting point is the infinite horizon macroeconomic model with banking and bank runs developed in Gertler and Kiyotaki (2014). In order to study recent financial booms and crises, in this chapter we disaggregate banking into wholesale and retail banks. Wholesale banks make loans to the non-financial sector funded primarily by borrowing from retail banks. The latter use deposits from households to make loans both to the non-financial sector and to the wholesale financial sector. Further, the size of the wholesale banking market arises endogenously. It depends on two key factors: (1) the relative advantage wholesale banks have in managing assets over retail banks; and (2) the relative advantage of retail banks over households in over-coming an agency friction that impedes lending to wholesale banks.

In the previous section we described the different layers of the wholesale sector, including origination, securitization and funding. For tractability, in our model we consolidate these various functions into a single type of wholesale bank. Overall, our model permits capturing financial stress in wholesale funding markets which was a key feature of the recent financial crisis.

There are three classes of agents: households, retail banks, and wholesale banks. There are two goods, a nondurable good and a durable asset, "capital." Capital does not depreciate and the total supply of capital stock is fixed at $\overline{K}$. Wholesale and retail banks use borrowed funds and their own equity to finance the acquisition of capital. Households lend to banks and also hold capital directly. Their respective total holdings of capital by each type of agent equals the total supply:

$$K_t^w + K_t^r + K_t^h = \overline{K}, \tag{1}$$

where K_t^w and K_t^r are the total capital held by wholesale and retail bankers and K_t^h is the amount held by households.

Agents of type j use capital and goods as inputs at t to produce output and capital at $t+1$, as follows:

$$\begin{array}{ccc} \textit{date } t & & \textit{date } t\textit{+1} \\ \left.\begin{array}{r} K_t^j \text{ capital} \\ F^j(K_t^j) \text{ goods} \end{array}\right\} & \rightarrow & \left\{\begin{array}{c} Z_{t+1}K_t^j \text{ output} \\ K_t^j \text{ capital} \end{array}\right. \end{array} \tag{2}$$

where type $j = w$, r and h stands for wholesale banks, retail banks, and households, respectively. Expenditure in terms of goods at date t reflects the management cost of screening and monitoring investment projects. In the case of retail banks, the management costs might also reflect various regulatory constraints. We suppose this management cost is increasing and convex in the total amount of capital, as given by the following quadratic formulation:

$$F^j(K_t^j) = \frac{\alpha^j}{2}(K_t^j)^2. \tag{3}$$

In addition we suppose the management cost is zero for wholesale banks and highest for households (holding constant the level of capital):

$$\alpha^w = 0 < \alpha^r < \alpha^h. \tag{Assumption 1}$$

This assumption implies that wholesale bankers have an advantage over the other agents in managing capital. Retail banks in turn have a comparative advantage over households. Finally, the convex cost implies that it is increasingly costly at the margin for retail banks and households to absorb capital directly. As we will see, this cost formulation provides a simple way to limit agents with wealth but lack of expertise from purchasing assets during a firesale.

In our decentralization of the economy, a representative household provides capital management services both for itself and for retail banks. For the latter, the household charges retail banks a competitive price f_t^r per unit of capital managed, where f_t^r corresponds to the marginal cost of providing the service:

$$f_t^r = F^{r\prime}(K_t^r) = \alpha^r K_t^r. \tag{4}$$

Households obtain the profit from this activity $f_t^r K_t^r - F^r(K_t^r)$.

3.2 Households

Each household consumes and saves. Households save either by lending funds to bankers or by holding capital directly in the competitive market. They may deposit funds in either retail or wholesale banks. In addition to the returns on portfolio investments, each household receives an endowment of nondurable goods, $Z_t W^h$, every period that varies proportionately with the aggregate productivity shock Z_t.

Deposits held in a bank from t to $t+1$ are one period bonds that promise to pay the non-contingent gross rate of return $\bar{R}_{t+1}$ in the absence of a bank run. In the event of a run depositors only receive a fraction x_{t+1} of the promised return on deposits, where x_{t+1} is the total liquidation value of bank assets per unit of promised deposit obligations. Accordingly, we can express the household's return on deposits, R_{t+1}, as follows:

$$R_{t+1} = \begin{cases} \overline{R}_{t+1} \text{ if no bank run} \\ x_{t+1}\overline{R}_{t+1} \text{ if run occurs} \end{cases} \tag{5}$$

where $0 \leq x_t < 1$. Note that if a run occurs all depositors receive the same pro rata share of liquidated assets.

For pedagogical purposes, we begin with a baseline model where bank runs are completely unanticipated events. Accordingly, in this instance the household chooses consumption and saving with the expectation that the realized return on deposits R_{t+1} equals the promised return $\overline{R}_{t+1}$ with certainty. In a subsequent section, we characterize the case where households anticipate that a bank run may occur with some likelihood.

Household utility U_t is given by

$$U_t = E_t \left(\sum_{i=0}^{\infty} \beta^i \ln C_{t+i}^h \right)$$

where C_t^h is household consumption and $0 < \beta < 1$. Let Q_t be the market price of capital. The household then chooses consumption, bank deposits D_t and direct capital holdings K_t^h to maximize expected utility subject to the budget constraint

$$C_t^h + D_t + Q_t K_t^h + F^h(K_t^h) = Z_t W^h + R_t D_{t-1} + (Z_t + Q_t) K_{t-1}^h + f_t^r K_t^r - F^r(K_t^r). \tag{6}$$

Here, consumption, saving and management costs are financed by the endowment, the returns on the saving, and the profit from providing management service to retail bankers.

Given that the household assigns a zero probability of a bank run, the first order conditions for deposits is given by

$$E_t(\Lambda_{t,t+1})R_{t+1} = 1 \tag{7}$$

where the stochastic discount factor $\Lambda_{t,\tau}$ satisfies

$$\Lambda_{t,\tau} = \beta^{\tau-t}\frac{C_t^h}{C_\tau^h}.$$

The first order condition for direct capital holdings is given by

$$E_t\left(\Lambda_{t,t+1}R_{kt+1}^h\right) = 1 \tag{8}$$

with

$$R_{kt+1}^h = \frac{Q_{t+1} + Z_{t+1}}{Q_t + F^{h\prime}(K_t^h)}$$

where $F'(K_t^h) = \alpha K_t^h$ and R_{t+1}^h is the household's gross marginal rate of return from direct capital holdings.

3.3 Banks

There are two types of bankers, retail and wholesale. Each type manages a financial intermediary. Bankers fund capital investments (which we will refer to as "business loans") by issuing deposits to households, borrowing from other banks in an interbank market and using their own equity, or net worth. Banks can also lend in the interbank market.

As we describe below, bankers may be vulnerable to runs in either the interbank market or the retail market, or both. In the former case, creditor banks suddenly decide to not rollover interbank loans. In the event of a run, the creditor banks receive a fraction x_{bt+1} of the promised return on the interbank credit, where x_{bt+1} is the total liquidation value of debtor bank assets per unit of debt obligations. Accordingly, we can express the creditor bank's return on interbank loans, R_{bt+1}, as follows:

$$R_{bt+1} = \begin{cases} \overline{R}_{bt+1} \text{ if no bank run} \\ x_{bt+1}\overline{R}_{bt+1} \text{ if run occurs} \end{cases} \tag{9}$$

where $0 \leq x_{bt} < 1$. If a run occurs all the creditor banks receive the same pro rata share of liquidated assets. As in the case of deposits, we continue to restrict attention to the case where banks runs are completely unanticipated, before turning in a subsequent section to the case of anticipated runs in wholesale funding markets.

Due to financial market frictions that we specify below, bankers may be constrained in their ability to raise external funds. To the extent they may be constrained, they will attempt to save their way out of the financing constraint by accumulating retained earnings in order to move toward one hundred percent equity financing. To limit this possibility, we assume that bankers have a finite expected lifetime: Specifically, each banker of type j (where $j = w$ and r for wholesale and retail bankers) has an i.i.d. probability σ^j of surviving until the next period and a probability $1-\sigma^j$ of exiting. This setup provides a simple way to motivate "dividend payouts" from the banking system in order to ensure that banks use leverage in equilibrium.

Every period new bankers of type j enter with an endowment w^j that is received only in the first period of life. This initial endowment may be thought of as the start up equity for the new banker. The number of entering bankers equals the number who exit, keeping the total constant.

We assume that bankers of either type are risk neutral and enjoy utility from consumption in the period they exit. The expected utility of a continuing banker at the end of period t is given by

$$V_t^j = E_t \left[\sum_{i=1}^{\infty} \beta^i (1-\sigma^j)(\sigma^j)^{i-1} c_{t+i}^j \right],$$

where $(1-\sigma^j)(\sigma^j)^{i-1}$ is probability of exiting at date $t+i$, and c_{t+i}^j is terminal consumption if the banker of type j exits at $t+i$.

The aggregate shock Z_t is realized at the start of t. Conditional on this shock, the net worth of "surviving" bankers j is the gross return on business loans net the cost of deposits and borrowing from the other banks, as follows:

$$n_t^j = (Q_t + Z_t)\, k_{t-1}^j - R_t d_{t-1}^j - R_{bt} b_{t-1}^j, \tag{10}$$

where d_{t-1}^j is deposit and b_{t-1}^j is interbank borrowing at $t-1$. Note that b_{t-1}^j is positive if bank j borrows and negative if j lends in the interbank market.

For new bankers at t, net worth simply equals the initial endowment:

$$n_t^j = w^j. \tag{11}$$

Meanwhile, exiting bankers no longer operate banks and simply use their net worth to consume:

$$c_t^j = n_t^j. \tag{12}$$

During each period t, a continuing bank j (either new or surviving) finances business loans $(Q_t + f_t^j)k_t^j$ with net worth, deposit and interbank debt as follows:

$$(Q_t + f_t^j)k_t^j = n_t^j + d_t^j + b_t^j, \tag{13}$$

where f_t^r is given by (4) and $f_t^w = 0$. We assume that banks can only accumulate net worth via retained earnings. While this assumption is a reasonable approximation of reality, we do not explicitly model the agency frictions that underpin it.

To motivate a limit on the bank's ability to raise funds, we introduce the following moral hazard problem: After raising fund and buying assets at the beginning of t, but still during the period, the banker decides whether to operate "honestly" or to divert assets for personal use. Operating honestly means holding assets until the payoffs are realized in period $t+1$ and then meeting obligations to depositors and interbank creditors. To divert means to secretly channeling funds away from investments in order to consumes personally.

The banker's ability to divert funds depends on the source: The banker can divert the fraction θ of total funds raised from households, where $0 < \theta < 1$. On the other hand, he/she can divert only the fraction $\theta\omega$ of funds financed by interbank borrowing, where $0 < \omega < 1$. Here we are capturing in a simple way that bankers lending in the wholesale market are more effective at monitoring the banks to which they lend than are households that supply deposits in the retail market. This efficiency advantage that banks have in lending to other banks provides an important reason for the growth of the wholesale market in our model, as we will show.

We assume that the process of diverting assets takes time: The banker cannot quickly liquidate a large amount assets without the transaction being noticed. For this reason the banker must decide whether to divert at t, prior to the realization of uncertainty at $t+1$. The cost to the banker of the diversion is that the creditors can force the intermediary into bankruptcy at the beginning of the next period.

The banker's decision at t boils down to comparing the franchise value of the bank V_t^j, which measures the present discounted value of future payouts from operating honestly, with the gain from diverting funds. In this regard, rational lenders will not supply funds to the banker if he has an incentive to cheat. Accordingly, any financial arrangement between the bank and its lenders must satisfy the following set of incentive constraints, which depend on whether the bank is a net borrower or lender in the interbank market:

$$V_t^j \geq \theta[n_t^j + d_t^j + \omega \cdot Max(b_t^j, 0)]. \tag{14}$$

As will become clear shortly, each incentive constraint embeds the constraint that the net worth n_t^j must be positive for the bank to operate: This is because the franchise value V_t^j will turn out to be proportional to n_t^j.

Given that bankers simply consume their net worth when they exit, we can restate the bank's franchise value recursively as the expected discounted value of the sum of net worth conditional on exiting and the value conditional on continuing as:

$$V_t^j = \beta E_t[(1-\sigma^j)n_{t+1}^j + \sigma^j V_{t+1}^j]. \tag{15}$$

The banker's optimization problem then is to choose $\left(k_t^j, d_t^j, b_t^j\right)$ each period to maximize the franchise value (15) subject to the incentive constraint (14) and the balance sheet constraints (10) and (13).

From the balance sheet constraints, we can express the growth rate of net worth of bank j as

$$\begin{aligned}
\frac{n_{t+1}^j}{n_t^j} &= \frac{Q_{t+1}+Z_{t+1}}{Q_t+f_t^j}\frac{\left(Q_t+f_t^j\right)k_t^j}{n_t^j} - R_{t+1}\frac{d_t^j}{n_t^j} - R_{bt+1}\frac{b_t^j}{n_t^j} \\
&= \left(R_{kt+1}^j - R_{bt+1}\right)\frac{\left(Q_t+f_t^j\right)k_t^j}{n_t^j} + \left(R_{bt+1}-R_{t+1}\right)\frac{d_t^j}{n_t^j} + R_{bt+1}
\end{aligned}$$

where R_{kt+1}^j is the rate of return on business loans for bank j:

$$R_{kt+1}^j = \frac{Q_{t+1}+Z_{t+1}}{Q_t+f_t^j}. \tag{16}$$

Because both the objective and constraints of the bank are constant returns to scale, its portfolio will be proportional to its size, as determined by its net worth n_t^j. Then the choice of bank j becomes

$$\begin{aligned}
\frac{V_t^j}{n_t^j} &= \underset{\frac{\left(Q_t+f_t^j\right)k_t^j}{n_t^j},\ \frac{d_t^j}{n_t^j}}{Max} \quad E_t\left[\beta\left(1-\sigma^j+\sigma^j\frac{V_{t+1}^j}{n_{t+1}^j}\right)\frac{n_{t+1}^j}{n_t^j}\right] \\
&= \underset{\frac{\left(Q_t+f_t^j\right)k_t^j}{n_t^j},\ \frac{d_t^j}{n_t^j}}{Max} \left[\mu_{kt}^j\frac{\left(Q_t+f_t^j\right)k_t^j}{n_t^j}+\mu_{bt}^j\frac{d_t^j}{n_t^j}+\nu_{bt}^j\right], \qquad (17)
\end{aligned}$$

subject to the incentive constraint

$$\frac{V_t^j}{n_t^j} \geq \theta\left[1+\frac{d_t^j}{n_t^j}+\omega\cdot Max\left(\frac{\left(Q_t+f_t^j\right)k_t^j}{n_t^j}-\frac{d_t^j}{n_t^j}-1,0\right)\right], \qquad (18)$$

where

$$\mu_{kt}^j = E_t\left[\beta\Omega_t^j\left(R_{kt+1}^j-R_{bt+1}\right)\right], \qquad (19)$$

$$\mu_{bt}^j = E_t\left[\beta\Omega_t^j\left(R_{bt+1}-R_{t+1}\right)\right], \qquad (20)$$

$$\nu_{bt}^j = E_t\left(\beta\Omega_t^j\right)R_{bt+1}, \qquad (21)$$

$$\Omega_t^j = 1-\sigma^j+\sigma^j\frac{V_{t+1}^j}{n_{t+1}^j}.$$

We can think of the ratio of franchise value to the net worth $\frac{V_t^j}{n_t^j}$ as the Tobin's Q. As will become clear, Tobin's Q may exceed unity due to the bank's financing constraint. The variable μ_{kt}^j is the discounted excess return on bank loans over interbank loans, μ_{bt}^j is the discounted excess cost of interbank loans relative to deposits, and ν_{bt}^j is the discounted marginal cost of an interbank loan. Observe also that the discount factor the bank uses to evaluate payoffs in $t+1$ is weighted by the multiplier Ω_t^j which is a probability weighted average of the marginal values of net worth to exiting and to continuing bankers at t+1. For an exiting banker at $t+1$ (which occurs with probability $1-\sigma^j$), the marginal value of an additional unit of net worth is simply unity, since he or she just consumes it. For a continuing banker (which occurs with probability σ^j), the marginal value equals Tobin's Q.

We defer the details of the formal bank maximization problems to Appendix A. Here we explain the decisions of wholesale and retail banks informally.

Because wholesale banks have a cost advantage over retail banks in making business loans, the rate of return on business loans is higher for the former than for the latter (see equation (16)). In turn, retail banks have an advantage over households in lending to wholesale banks due to their relative advantage in recovering assets in default. Therefore, in equilibrium, wholesale banks borrow from retail banks in the interbank market to make business loans. Indeed the only reason retail banks directly make business loans is because wholesale banks may be constrained in the amount of business loans they can make.

Suppose that the excess return on business loans over inter-bank borrowing is positive, i.e. $0 < \mu_{kt}^w$.[13] Then wholesale banks will want to expand the inter-bank borrowing and business loan. The expansion is limited only because the incentive constraint binds.[14] Accordingly combining the Bellman equation (17) and the incentive constraint (18) with equality yields expressions for the bank's ratio of business loans to net worth when it is constrained and for it's Tobin's Q value that arises in this situation:

$$\frac{Q_t k_t^w}{n_t^w} = \frac{1}{\theta\omega - \mu_{kt}^w}\left\{\nu_{bt}^w - \theta(1-\omega) - [\theta(1-\omega) - \mu_{bt}^w]\frac{d_t^w}{n_t^w}\right\}, \tag{22}$$

$$\frac{V_t^w}{n_t^w} = \frac{\theta}{\theta\omega - \mu_{kt}^w}\left\{\omega\nu_{bt}^w - (1-\omega)\mu_{kt}^w + [\omega\mu_{bt}^w - (1-\omega)\mu_{kt}^w]\frac{d_t^w}{n_t^w}\right\}. \tag{23}$$

To determine whether wholesale banks issue deposits, we maximize (23) with respect to d_t^w. The choice of deposits by wholesale banks is then given by

$$\begin{aligned} d_t^w &= 0, \text{ if } \omega\mu_{bt}^w < (1-\omega)\mu_{kt}^w, \\ \omega\mu_{bt}^w &= (1-\omega)\mu_{kt}^w, \text{ if } d_t^w > 0. \end{aligned} \tag{24}$$

The wholesale bank faces the following trade-off in using retail deposits: If the deposit rate is less than the interbank rate so that $\mu_{bt}^w > 0$, then the bank gains from issuing deposits to reduce interbank borrowing. On the other hand, because households are less efficient in monitoring wholesale bank behavior, they will apply a tighter limit on the amount they are willing

[13] We will later choose parameters so that $0 < \mu_{kt}^w$ holds in the neighborhood of the deterministic steady state, and numerically show that this inequality always holds in our dynamic equilibrium.

[14] We can prove that $\mu_{kt}^w > 0$ implies $\mu_{kt}^w < \theta\omega$ in equilibrium. See Appendix A.

to lend than will retail banks. If $\omega\mu_{bt}^{w} < (1-\omega)\mu_{kt}^{w}$, the cost exceeds the benefit. In this instance the wholesale bank does not use retail deposits, relying entirely on interbank borrowing for external finance. Everything else equal, by not issuing retail deposits, the wholesale bank is able to raise its overall leverage in order to make more business loans relative to its equity base. This incentive consideration accounts for why the wholesale bank may prefer interbank borrowing to issuing deposits, even if the interbank rate lies above the deposit rate.[15]

Combining (22), (23), and (24) yields:

$$\frac{Q_t k_t^w}{n_t^w} + \frac{1-\omega}{\omega}\frac{d_t^w}{n_t^w} = \frac{\nu_{bt}^w - \theta(1-\omega)}{\theta\omega - \mu_{kt}^w} \equiv \phi_t^w, \tag{25}$$

$$\frac{V_t^w}{n_t^w} = \frac{\theta\left[\omega\nu_{bt}^w - (1-\omega)\mu_{kt}^w\right]}{\theta\omega - \mu_{kt}^w}. \tag{26}$$

The right hand side of (25) is the maximum feasible value of the bank's ratio of assets to net worth that satisfies the incentive constraint. We refer to this value, which we define as ϕ_t^w, as the maximum feasible "leverage multiple" that the wholesale bank can maintain. It is an increasing function of the discounted excess return on business loans over interbank loans (μ_{kt}^w) and a decreasing function of the interbank asset diversion rate $\theta\omega$. Note that given its equity base, the wholesale bank achieves maximum feasible amount of business lending by reducing retail deposits d_t^w to zero.

We next turn to the retail banker's problem. Because the retail banks lend in the interbank market, we know $\left(Q_t + f_t^r\right)k_t^r < d_t^r + n_t^r$. Then from from the Bellman equation and the incentive constraint $(17, 18)$, we obtain

$$\begin{aligned} \mu_{kt}^r &\leq 0. \\ \mu_{kt}^r &= 0, \text{ if } k_t^r > 0. \end{aligned} \tag{27}$$

Equation (27) requires that the discounted excess return on business loans over interbank loans must zero, in order for retail banks to make both types

[15]Under our baseline parametrization, wholesale banks borrow exclusively from retail banks. We view this as the case the that best corresponds to the wholesale banking system on the eve of the Great Recession. Circumstances do exist where wholesale banks will borrow from households as well as retail banks. One might interpret his situation as corresponding to the consolidation of wholesale and retail bank in the wake of the crisis, or perhaps the period before the rapid growth of wholesale banking when retail banks were performing many of the same activities as we often observe in continental Europe and Japan.

of loans. We can show that, as long as wholesale banks are financially constrained in the amount of business loans they make, retail banks and households take some of this business. Even though they are less efficient than wholesale banks, the limits to wholesale bank arbitrage leads to excess returns on some business loans that make them sufficiently profitable for retail banks and households to hold.

As with wholesale banks, we will choose a parametrization where the incentive constraint binds, (which implies that retail banks make a positive excess return on making interbank loans, i.e. $0 < \mu_{bt}^r$). Then from (17) and (18) with equality, we get the ratio of business and interbank loans to the net worth, as follows.

$$\frac{(Q_t + f_t^r)\, k_t^r}{n_t^r} + \frac{-b_t^r}{n_t^r} = \frac{d_t^r}{n_t^r} + 1 = \frac{\nu_{bt}^r - \mu_{bt}^r}{\theta - \mu_{bt}^r} \equiv \phi_t^r. \tag{28}$$

The right hand side of (28) is the maximum feasible leverage multiple of the retail bank, i.e. the maximum value that satisfies the incentive constraint. Given $\nu_{bt}^r > \theta$, the leverage multiple is an increasing function of discounted excess cost of interbank loans over deposits and a decreasing function of the asset diversion θ. Tobin's Q for the retail bank then becomes

$$\frac{V_t^r}{n_t^r} = \frac{\theta(\nu_{bt}^r - \mu_{bt}^r)}{\theta - \mu_{bt}^r}. \tag{29}$$

3.4 Aggregation and Equilibrium without Bank Runs

Given that the ratio of assets and liabilities to net worth is independent of individual bank-specific factors and given a parametrization where in equilibrium the incentive constraints are binding, we can aggregate across banks to obtain relations between total assets and net worth for both the wholesale and retail banking sectors. Let $Q_t K_t^w$ and $Q_t K_t^r$ be total business loans held by wholesale and retails banks, D_t^w and D_t^r be their deposits, B_t be total interbank debt, and N_t^w and N_t^r total net worth in each respective banking sector. Then we have:

$$Q_t K_t^w + \frac{1-\omega}{\omega} D_t^w = \phi_t^w N_t^w, \tag{30}$$

$$(Q_t + f_t^r) K_t^r + B_t = \phi_t^r N_t^r, \tag{31}$$

with

$$Q_t K_t^w = N_t^w + D_t^w + B_t, \tag{32}$$

$$(Q_t + f_t^r) K_t^r + B_t = D_t^r + N_t^r, \tag{33}$$

and

$$\begin{aligned} D_t^w &= 0, \text{ if } \omega \mu_{bt}^w < (1-\omega)\mu_{kt}^w, \\ \omega \mu_{bt}^w &= (1-\omega)\mu_{kt}^w, \text{ if } D_t^w > 0. \end{aligned} \tag{34}$$

Summing across both surviving and entering bankers yields the following expression for the evolution of N_t :

$$N_t^w = \sigma^w[(Z_t + Q_t)K_{t-1}^w - R_t D_{t-1}^w - R_{bt} B_{t-1}] + W^w, \tag{35}$$

$$N_t^r = \sigma^r[(Z_t + Q_t)K_{t-1}^r - R_t D_{t-1}^r + R_{bt} B_{t-1}] + W^r, \tag{36}$$

where $W^j = (1-\sigma^j)w^j$ is the total endowment of entering bankers. The first term is the accumulated net worth of bankers that operated at $t-1$ and survived to t, which is equal to the product of the survival rate σ^j and the net earnings on bank assets.

Total bankers consumption equals the sum of the net worth of exiting bankers in each sector:

$$\begin{aligned} C_t^b = \; & (1-\sigma^w)[(Z_t + Q_t)K_{t-1}^w - R_t D_{t-1}^w - R_{bt} B_{t-1}] \\ & +(1-\sigma^r)[(Z_t + Q_t)K_{t-1}^r - R_t D_{t-1}^r + R_{bt} B_{t-1}]. \end{aligned} \tag{37}$$

Total output Y_t is the sum of output from capital, household endowment $Z_t W^h$ and bank endowment W^r and W^i :

$$Y_t = Z_t + Z_t W^h + W^r + W^i.$$

Finally, output is either used for management costs, or consumed by households and bankers:

$$Y_t = F^h(K_t^h) + F^r(K_t^r) + C_t^h + C_t^b. \tag{38}$$

The recursive competitive equilibrium without bank runs consists of aggregate quantities $\left(K_t^w, K_t^r, K_t^h, B_t, D_t^w, D_t^r, N_t^w, N_t^r, C_t^b, C_t^h\right)$, prices $(Q_t, R_{t+1}, R_{bt+1}, f_t^r)$ and bankers' franchise values and leverage multiples $\left(\mu_{kt}^j, \mu_{kt}^j, \nu_{bt}^j, \frac{V_t^j}{n_t^j}, \phi_t^j\right)_{j=w,r}$ as a function of the state variables $\left(K_{t-1}^w, K_{t-1}^r, R_{bt} B_{t-1}, R_t D_{t-1}^w, R_t D_{t-1}^r, Z_t\right)$, which satisfy equations $(1, 4, 7, 8, 19-38)$.[16]

[16] In total we have a system of 24 equations. Notice that $(19-21)$ have two equations.

3.5 Unanticipated Bank Runs

In this section we consider unanticipated bank runs. We defer an analysis of anticipated bank runs to Section 5. In general three types of runs are possible: (i) a run on wholesale banks leaving retail banks intact; (ii) a run on just retail banks; and (iii) a run on both the wholesale and retail bank sectors. We focus mainly on (i) because it corresponds most closely to what happened in practice, but we also briefly discuss the other two cases. We also restrict the attention to the case in which wholesale banks entirely rely on interbank borrowing for external finance.

3.5.1 Conditions for a Wholesale Bank Run Equilibrium

The runs we consider are runs on the entire wholesale banking system, not on individual wholesale banks. Indeed, so long as an asset firesale by an individual bank is not large enough to affect asset prices, it is only runs on the system that will be disruptive. Given the homogeneity of wholesale banks in our model, the conditions for a run on the wholesale banking system will apply to each individual wholesale bank.

As we noted earlier, what we have in mind for a run is a spontaneous failure of the bank's creditors to roll over their short term loans. In particular, at the beginning of period t, before the realization of returns on bank assets, retail banks lending to a wholesale bank decide whether to roll over their loans with the bank. If they choose to "run", the wholesale bank liquidates its capital and turns the proceeds over to its retail bank creditors who then either acquire the capital or sell it to households. Importantly, both the retail banks and households cannot seamlessly acquire the capital being liquidated in the firesale by wholesale banks. The retail banks face a capital constraint which limits asset acquisition and are also less efficient at managing the capital than are wholesale banks. Households can only hold the capital directly and are even less efficient than retail banks in doing so. Let Q_t^* be the price of capital in the event of a forced liquidation of the banking system. Then a run on the entire wholesale bank sector is possible if the liquidation value of bank assets $(Z_t + Q_t^*)K_{t-1}^w$ is smaller than its outstanding liability to the interbank creditors, $R_{bt}B_{t-1}$, in which case the bank's net worth would be wiped out. In this instance the recovery rate in the event of a bank run, x_{bt}^w,

By Walras' law, the household budget constraint (6) is satisfied as long as deposit market clears as $D_t = D_t^w + D_t^r$.

is the ratio of $(Z_t + Q_t^*)K_{t-1}^w$ to $R_{bt}B_{t-1}$. Then the condition for a bank run equilibrium to exist is that the recovery rate is less than unity, i.e.

$$x_{bt}^w = \frac{(Q_t^* + Z_t)K_{t-1}^w}{R_{bt}B_{t-1}} < 1. \tag{39}$$

The condition determining the possibility of a bank run depends on two key endogenous factors, the liquidation price of capital Q_t^* and the condition of bank balance sheets. From (35), we can obtain a simple condition for a bank run equilibrium in terms of just three variables:

$$x_{bt}^w = \frac{R_{kt}^{w*}}{R_{bt}} \cdot \frac{\phi_{t-1}^w}{\phi_{t-1}^w - 1} < 1 \tag{40}$$

with

$$R_{kt}^{w*} \equiv \frac{Z_t + Q_t^*}{Q_{t-1}},$$

where R_{kt}^{w*} is the return on bank assets conditional on a run at t, and ϕ_{t-1}^w is the leverage multiple of wholesale bank at $t-1$. A bank run equilibrium exists if the realized rate of return on bank assets conditional on liquidation of assets R_{kt}^{w*} is sufficiently low relative to the gross interest rate on interbank loans, R_{bt}, and the leverage multiple is sufficiently high to satisfy condition (40). Note that the expression $\frac{\phi_{t-1}^w}{\phi_{t-1}^w-1}$ is the ratio of bank assets $Q_{t-1}K_{t-1}^w$ to deposits B_{t-1}, which is decreasing in the leverage multiple. Also note that the condition for a run does not depend on individual bank-specific factors since $(R_{kt}^{w*}/R_{bt},\, \phi_{t-1}^w)$ are the same for all in equilibrium.

Since R_{kt}^{w*}, R_{bt} and ϕ_{t-1}^w are all endogenous variables, the possibility of a bank run may vary with macroeconomic conditions. The equilibrium absent bank runs (that we described earlier) determines the behavior of R_{bt} and ϕ_{t-1}^w. The value of R_{kt}^{w*} is increasing in the liquidation price Q_t^*, which depends on the behavior of the economy, as we show in the next sub-section.

3.5.2 The Liquidation Price

To determine Q_t^* we proceed as follows. A run by interbank creditors at t induces all wholesale banks that carried assets from $t-1$ to fully liquidate their asset positions and go out of business.[17] Accordingly they sell all their

[17]See Uhlig (2010) for an alternative bank run model with endogenous liquidation prices.

assets to retail banks and households, who hold them at t. The wholesale banking system then re-builds itself over time as new banks enter. For the asset firesale during the panic run to be quantitatively significant, we need there to be at least a modest delay in the ability of new banks to begin operating. Accordingly, we suppose that new wholesale banks cannot begin operating until the period after the panic run. Suppose for example that during the run it is not possible for retail banks to identify new wholesale banks that are financially independent of the wholesale banks being run on. New wholesale banks accordingly wait for the dust to settle and then begin raising fund in the interbank market in the subsequent period. The results are robust to alternative timing assumptions about the entry of new banks.

Accordingly, when wholesale banks liquidate, they sell all their assets to retail banks and households in the wake of the run at date t, implying

$$\overline{K} = K_t^r + K_t^h. \tag{41}$$

The wholesale banking system then rebuilds its equity and assets as new banks enter at t+1 onwards. Given our timing assumptions and Equation (35), bank net worth evolves in the periods after the run according to

$$\begin{aligned} N_{t+1}^w &= (1+\sigma^w)W^w, \\ N_{t+i}^w &= \sigma^w[(Z_{t+i}+Q_{t+i})K_{t+i-1}^w - R_{bt+i}B_{t+i-1}] + W^w, \text{ for all } i \geq 2. \end{aligned}$$

Rearranging the Euler equation for the household's capital holding (8) yields the following expression for the liquidation price in terms of discounted dividends Z_{t+i} net the marginal management cost $\alpha^h K_{t+i}^h$.

$$Q_t^* = E_t \left[\sum_{i=1}^{\infty} \Lambda_{t,t+i}(Z_{t+i} - \alpha^h K_{t+i}^h) \right] - \alpha^h K_t^h. \tag{42}$$

Everything else equal, the longer it takes for the banking sector to recapitalize (measured by the time it takes K_{t+i}^h to fall back to steady state), the lower will be the liquidation price. Note also that Q_t^* will vary with cyclical conditions. In particular, a negative shock to Z_t will reduce Q_t^*, possibly moving the economy into a regime where bank runs are possible.

4 Numerical Experiments

In this section we examine how the long-run properties of the model can account for the growth of the wholesale banking sector and then turn to studying the cyclical responses to variations in productivity and runs. Overall these numerical examples provide a description of the tradeoff between growth and stability associated with an expansion of the shadow banking sector and illustrate the real effects of bank runs in our model.

4.1 Calibration

Here we describe our baseline calibration. This is meant to capture the state of the economy at the onset of the financial crisis in 2007.

There are 13 parameters in the model:

$$\left\{\beta, \alpha^h, W^h, \sigma^r, \alpha^r, W^r, \theta, \sigma^w, \alpha^w, W^w, \omega, \sigma_z, \rho_z\right\}.$$

Their values are reported in Table 1, while Table 2 shows the Steady State values of the equilibrium allocation.

We take the time interval in the model to be a quarter. We use conventional values for households' discount factor, $\beta = .99$, and the parameters governing the stochastic process for dividends, $\sigma_z = .05$ and $\rho_z = .9$. We set W^h so that households endowment income is twice as big as the aggregate capital income.

We calibrate managerial costs of intermediating capital for households and retail bankers, α^h and α^r, in order to obtain the spread between deposit and interbank interest rates as well as the spread between interbank and business loan rates both to be 1.2% in annual in steady state.

The fraction of divertible assets purchased by raising deposits, θ, and interbank loans, $\omega\theta$, are set in order to get leverage ratios for retail bankers and wholesale bankers of 10 and 25 respectively.

Our retail banking sector comprises of commercial banks, open end Mutual Funds and Money Market Mutual Funds (MMMF). In the case of Mutual Funds and MMMF the computation of leverage is complicated by the peculiar legal and economic details of the relationship between these institutions, their outside investors and sponsors.[18] Hence, our choice of 10 quite closely reflects the actual leverage ratios of commercial banks, which is the only sector for which a direct empirical counterpart of leverage can be easily computed.

To set our target for wholesale leverage we decide to focus on private institutions within the wholesale banking sector that relied mostly on short

[18] Here we could cite Cecilia Parlatore's thesis and McCabe's paper.

term debt. A reasonable range for the leverage multiple for such institutions goes from around 10 for some ABCP issuers [19] to values of around 40 for brokers dealers in 2007. Our choice of 25 is the average of these two extreme values.

The survival rates of wholesale and retail bankers, σ^w and σ^r, are set in order for the distribution of assets across sectors to match the actual distribution in 2007. Finally, we set W^r to make new entrants net worth being equal to 1% of total retail banks net worth and W^w to ensure that wholesale bankers are perfectly specialized.

4.2 Long Run Effects of Financial Innovation

As mentioned in Section 2, the role of wholesale banks in financial intermediation has grown steadily from the 1980's to the onset of the financial crisis. This growth was largely accomplished through a series of financial innovations that enhanced the borrowing capacity of the system by relying on securitization to attract funds from institutional investors. While our model abstracts from the details of the securitization process, we capture its direct effects on wholesale banks' ability of raising funds in interbank markets with a reduction in the severity of the agency friction between retail banks and wholesale banks, which is captured by parameter ω. Hence, in this section we study the long run behavior of financial intermediation in response to a decrease in ω and compare it to the low frequency dynamics in financial intermediation documented in Section 2.

Figure 9 shows how some key variables depend upon ω in the steady state. The direct effect of ameliorating the agency problem between wholesale and retail banks is a relaxation of wholesale banks' incentive constraints. The improved ability of retail banks to seize the assets of wholesale bankers in the case of cheating allows wholesale bankers to borrow more aggressively from retail bankers. This can be seen in the maximum leverage multiple of the wholesale banks in the steady state, which is,

$$\phi^w(\omega) = 1 + \frac{1}{\theta\omega}\left[\frac{(1-\sigma^w)}{\sigma^w}\frac{\beta\left(1-\frac{W^w}{N^w}\right)}{1-\beta\left(1-\frac{W^w}{N^w}\right)} - \theta\right].$$

[19] The same caveat as in the case of MMFs applies here because it is very complicated to factor in the various lines of credit that were provided by the sponsors of these programs.

Notice that a decreases in ω leads to higher wholesale leverage for each fixed level of net worth of wholesale banks (given the term in brackets is positive).

The general equilibrium effects of a lower ω work through various channels. For an economy with a lower interbank friction ω, the leverage multiple of the wholesale banking sector is higher, with a larger capital K^w and a larger amount interbank borrowing B by wholesale banking sector. Conversely, capital intermediated by retail banks K^r and households K^h tends to be lower. In the absence of bank runs, the relative shift of assets to the wholesale banking sector implies a more efficient allocation of capital and consequently a higher capital price Q_t. The flow of assets into wholesale banking, further, reduces the spread between the return on capital for wholesale banks and the interbank rate, as well as the spread between interbank and deposit rates. Despite lower spreads, both wholesale and retail banks enjoy higher franchise values thanks to the positive effect of higher leverage on total returns on equity. A unique aspect of financial innovation due to a lower friction in the interbank market is that the borrowing and lending among banks tends to be larger relative to the flow-of-funds from ultimate lenders (households) to ultimate borrowers (nonfinancial business). (See Appendix B).

Figure 10 and Figure 11 compare the steady state effect of financial innovations on some key measures of financial intermediation with the observed low frequency trends in their empirical counterparts. In particular, we assume that the value of ω in our baseline calibration results from a sequence of financial innovations that took place gradually from the 1980's to the financial crisis. For simplicity, we divide our sample into 4 periods of equal length and assign a value of ω to each subsample in order to match the observed percentage of intermediation of wholesale bankers over the period. In order to compute leverage of wholesale banks in Figure 11, we compute leverage of the three sectors within the wholesale banking sector that were mainly responsible for the growth of wholesale intermediation. Overall, the steady state comparative statics capture quite well the actual low frequency dynamics in financial intermediation observed over the past few decades.[20]

[20] The model overstatement of the role of retail intermediation relative to household direct holding of assets can be rationalized by the lack of heterogeneity in ultimate borrowers' funding sources since, in the data, households mainly hold equities while intermediaries are responsible for most debt intermediation. Introducing a different type of asset for which intermediaries have a smaller efficiency would then help to reconcile the evolution of the distribution of capital across sectors predicted by the model in response to financial

4.3 Recessions and Runs

Figure 12 shows the response of the economy to an unanticipated negative five percent shock to productivity Z_t, assuming that a run does not happen. To capture the effects of financial liberalization on the cyclical properties of the economy, we consider both our baseline parameterization and one with a higher ω which we set to be equal to the one associated with the early 1980's in Figure 10. In both cases the presence of financial constraints activates the familiar financial accelerator mechanism of Bernanke and Gertler (1989) and Kiyotaki and Moore (1997). Leverage amplifies the effects of the drop in Z_t on bankers' net worth, inducing a tightening of financial constraints; this reduces asset prices and feeds back into lower net worth.

Higher exposure to variations in Z_t and higher leverage make this effect stronger for wholesale banks that are forced into a firesale liquidation of their assets, which in turn leads them to reduce their demand for interbank loans. As a result, retail bankers increase their asset holdings and absorb, together with households, the capital flowing out of the wholesale banking sector. However, the relative inefficiency of these agents in intermediating assets makes this process costly as shown by the rise in the cost of bank credit and the amplification in the drop in output. Under our baseline calibration, spreads between gross borrowing costs for non financial borrowers and the risk free rate increase by forty basis points and output drops by seven percent.

Notice that financial liberalization induces a reallocation of risk from retail bankers to wholesale bankers. The sensitivity of net worth, leverage and spreads to variations in Z_t increases for wholesale and decreases for retail banks after financial liberalization. The overall financial accelerator is smaller in the economy after financial innovation which features a smaller increase in the total spread between the rate of return on capital (to ultimate borrowers) and deposit rate (to ultimate lenders) and a smaller drop in asset prices following the drop in Z_t. With a lower financial friction in the interbank market after financial innovation, the wholesale and retail banks are more financially integrated, and the leverage multiple of the entire banking system (the ratio of aggregate finance for nonfinancial sector to the net worth of banks) is significantly lower than the leverage multiple of individual banks, due to the large interbank positions. Hence, focusing on the no run equilibrium, financial innovation has beneficial effects both on the long run and on the cyclical properties of the economy. This conclusion is changed

innovation with the empirical one.

once we consider bank runs.

Figure 13 and Figure 14 describe the effects of bank runs. In particular we assume that two periods after the unanticipated drop in Z_t, retail investors stop rolling over short term debt issued by wholesale banks, inducing them to liquidate all of their assets and go bankrupt. Although in practice retail bankers were shielded from runs thanks to deposit insurance, for illustrative purposes Figure 13 considers the case in which the run on wholesale banks triggers a run on retail banks as well, while Figure 14 describes the more natural case in which households keep their deposits in retail banks even after a run on wholesale banks.

In both cases we plot a variable that indicates at each time t whether a run is possible at time $t+1$. To construct this variable we define

$$Run_t^{all} = \min\left\{1 - x_{b,t}, 1 - x_t\right\}$$

$$Run_t^{w} = 1 - x_{bt}$$

where x_{bt} and x_t are the recovery rates on wholesale and retail debt respectively. Hence, in order for a run of a given type to exist the associated run variable must be positive.

As shown by the Run^w variable in Figure 14, a run on wholesale banks is not possible in the steady state under both parametrizations considered. However, while in the economy with tighter financial constraints a run is still not possible even after a five percent drop in Z_t, the same drop in productivity is big enough to make a run on wholesale banking possible in the economy with lower ω. This is because the larger share of wholesale intermediation makes liquidation more costly in the economy after financial innovation and high leverage of wholesale banks increases the amplification of asset price reductions on net worth losses.

As explained in Section 3.5.1, the run on wholesale banks forces them into bankruptcy and results in K^w dropping to 0. Households and retail banks are forced to absorb all of the wholesale banks' assets, inducing asset prices to drop by about 6% in total. The intermediation costs associated with the reallocation of assets to less efficient agents leads to an additional contraction of output of around 6%, resulting in an overall drop of about 12%.

As new wholesale bankers resume operations from the period after the run, high levels of spreads for both retail and wholesale bankers allows them to increase their leverage and recapitalize financial intermediaries thanks to above average retained earnings. The reintermediation process however is rather lengthy and output remains depressed for a prolonged period of time.

5 Anticipated Runs

So far, we have focused on the case in which runs are completely unexpected. In this section we study how the equilibrium changes if agents anticipate that a run will occur with positive probability in the future, focusing on the more realistic case of a run on wholesale bankers only. The Appendix contains a detailed description of the equilibrium in this case. Here we describe the key forces through which anticipation of a run in the future affects financial intermediation.

The main difference from the unanticipated case is in the market for interbank loans. In particular, once runs are anticipated, retail bankers internalize how wholesale bankers' leverage affects returns on interbank loans in case of a run and adjust the required promised rate $\bar{R}_{bt+1}$ accordingly. We denote by p_t the time t probability that retail banks will run on wholesale banks at time $t+1$.[21] The first order condition determining retail bankers supply of loans becomes:[22]

$$E_t\left\{(1-p_t)\,\Omega^r_{t+1}\left[\left(R^r_{kt+1}-\bar{R}_{bt+1}\right)\right]+p_t\Omega^{r*}_{t+1}\left[\left(R^{r*}_{kt+1}-x_{bt+1}\bar{R}_{bt+1}\right)\right]\right\}=0 \tag{43}$$

where

$$\Omega^{r*}_{t+1}=1-\sigma+\sigma\frac{V^{r*}_{t+1}}{n^{r*}_{t+1}}$$

is the value of retail bankers' net worth if a run occurs at $t+1$. Using equation (40) to substitute for x_{bt+1} in (43) we obtain a menu of promised rates:[23]

$$\bar{R}^b_{t+1}\left(\phi^w_t\right)=\frac{E_t\left[(1-p_t)\,\Omega^r_{t+1}R^r_{kt+1}+p_t\Omega^{r*}_{t+1}\left(R^{r*}_{kt+1}-R^{w*}_{kt+1}\frac{\phi^w_t}{\phi^w_t-1}\right)\right]}{(1-p_t)\,E_t(\Omega^r_{t+1})} \tag{44}$$

Notice that $\bar{R}_{bt+1}\left(\phi^w_t\right)$ is an increasing function ϕ^w_t. This is because as leverage increases, retail bankers suffer larger losses on interbank loans if a run occurs, i.e., x^w_{bt+1} decreases. This induces them to require higher returns in the event of no run, to compensate for the higher losses in the event of a run.

[21] The determination of this probability of "observing a sunspot" will be discussed below.

[22] Consistent with the notation above, the expectation operator here is only over the uncertainty with respect to Z_t while the uncertainty arising from sunspots is explicitly accounted for.

[23] This is the relevant function for values of leverage high enough to induce bankruptcy in case of a run.

When choosing their portfolios, wholesale bankers will now have to factor in that changes in their leverage affect their cost of credit according to Equation (44) . This preserves homogeneity of the problem but the franchise value of the firm will change to reflect that with probability p_t the bank will be forced to liquidate assets at price Q^*_{t+1} in the subsequent period. This will have the effect of reducing the franchise value of wholesale banks, hence tightening their financial constraints.

In particular the franchise value of a perfectly specialized wholesale bank will still be given by

$$\frac{V^w_t}{n^w_t} = \mu^w_{kt}\frac{Q_t k^w_t}{n^w_t} + \nu^w_{bt},$$

but the discounted excess return on wholesale investment and the discounted marginal cost of interbank loans are now[24]

$$\mu^w_{kt} = \beta E_t \left\{ \Omega^w_{t+1} \left[(1-p_t)(R^w_{kt+1} - R^o_{bt+1}) + p_t \frac{R^{w*}_{kt+1} - R^{r*}_{kt+1}}{E_t(\Omega^r_{t+1})} \right] \right\}$$

$$\nu^w_{bt} = \beta E_t(\Omega^w_{t+1}) \left[(1-p_t)R^o_{bt+1} + p_t \frac{E_t(\Omega^{r*}_{t+1} R^{r*}_{kt+1})}{E_t(\Omega^r_{t+1})} \right]$$

where $R^o_{bt+1} = \frac{E_t(\Omega^r_{t+1} R^r_{kt+1})}{E_t(\Omega^r_{t+1})}$ is the riskless interbank rate conditional on no bank run.

In order to pin down a state dependent probability of a run, we follow Gertler and Kiyotaki (2015). In particular we assume that at each time t the probability of transitioning to a state where a run on wholesale banks occurs is given by a reduced form decreasing function of the expected recovery rate $E_t x^w_{bt+1}$ as follows,

$$p_t = \left[1 - E_t(x^w_{bt+1})\right]^{\delta}. \tag{45}$$

Although we don't endogenize the functional dependence of p_t from the state of the economy, the above formulation allows us to capture the idea that as wholesale balance sheet positions weaken, the likelihood of a run increases. This same qualitative conclusion would follow, for example, if the probability of a run was determined endogenously by introducing imperfect

[24]Here we are already assuming that wholesale bankers will choose a leverage high enough to result in bankruptcy when a run occurs. See the Appendix for a detailed description of hte wholesale banker's problem when runs are anticipated. There, we derive the conditions that ensure that it is optimal for wholesale bankers to default in the event of a run.

information, as in the global games approach developed by Morris and Shin (1998).

Figure 15 demonstrates how anticipation effects work to increase financial amplification of shocks in the model. The solid line is the response of the economy to an unanticipated five percent shock to Z_t when agents anticipate that a run can happen at each time $t+1$ with probability p_t as determined in Equation (45).[25] To isolate the effect of the anticipation of the run, we suppose in this case that the run never actually occurs ex-post. For comparison, the dotted line reports the responses of the baseline economy in which individuals assign probability zero to a bank run.

While it is still the case that in steady state a run cannot occur, the shock to Z_t leads the probability of a run to increase to 9%. As wholesale bankers' balance sheets weaken and the liquidation price decreases, retail bankers expect more losses on interbank loans in case of a run and the probability of coordinating on a run equilibrium increases as a result. The increase in p_t leads to a sharp contraction in the supply of interbank credit and a further tightening of wholesale bankers financial constraints. This, in turn, results in an overall reduction in their net worth of about 80% compared to a 60% in the baseline and to a spike in spreads between business loan and interbank loan rates that increases the spread by 180 basis points compared to only 30 in the baseline. As wholesale banks are forced to downsize their operations, total interbank credit falls by about 60%, more than twice the percentage drop in the baseline. These massive withdrawals of funds from wholesale markets is the model counterpart to the "slow runs" on the ABCP market in 2007. These disruptions in wholesale funding markets are then transmitted to the rest of the economy inducing a drop in asset prices of four percent and a total contraction of output of ten percent.

Figure 16 shows the case in which the run actually occurs two periods after the realization of the shock to Z_t. There are two main differences with respect to the analogous experiment performed in the case of unanticipated runs depicted in Figure 14. First, the initial increase in the probability of a run that precedes the actual run allows the model to capture the "slow runs" followed by "fast runs" in wholesale funding markets that was a central feature of the financial crisis, as discussed in the Introduction. Second, the run induces a further increase in the probability of additional runs in the future, that goes back to about 9% the period after the run occurs. This ham-

[25] In the numerical simulations below we pick δ to be $\frac{1}{2}$.

pers wholesale bankers ability to increase their leverage and generates higher spreads in the interbank market preventing the relatively smooth increase in asset prices that characterizes the recovery in the baseline model.

Figure 17 shows how the model with anticipated runs can reproduce some key features of the financial disruptions that occurred in 2007 and 2008. In particular, in the top two panels we compare the model predicted path for interbank spreads, $\bar{R}^b_{t+1} - R_{t+1}$, and excess finance premium, $ER^w_{k,t+1} - R_{t+1}$, with their empirical counterparts over the period going from 2007Q2 to 2009Q4. For the interbank spreads we choose the ABCP spread, since the first "slow runs" in wholesale funding markets in the third quarter of 2007 took place in the ABCP market. The measure of excess borrowing costs is the Excess Bond Premium of Gilchrist and Zakrajsek (2012). The bottom panel plots the model implied evolution of bank equity, measured by $V^w_t + V^r_t$, against the S&P financial index. We assume that the economy is in steady state in 2007Q2 and the unanticipated shock hits in 2007Q3 followed by a run on wholesale banks in 2008Q2. In the data excess borrowing costs lag financial spreads, so the model predicts a stronger initial increase in $ER^w_{k,t+1} - R_{t+1}$ and attributes a slightly smaller proportion of the increase to interbank spreads, probably due to the behavior of the risk free rate. On the other hand, the faster decline in spreads in the data after 2009 can be attributed to the effects of government intervention in this period. Finally, the higher persistence of spreads help explain why the franchise value of banks in the model is flatter after 2009. Overall, the experiment can capture the credit spreads and bank equity dynamics reasonably well.

6 Appendix

6.1 Appendix A: Detail of Bank Choice

The wholesale bank chooses $\left(\frac{Q_t k^w_t}{n^w_t}, \frac{d^w_t}{n^w_t}\right)$ to maximize Tobin's Q (17) subject to the incentive constraint (18). Letting λ^w_t be Lagrangian multiplier of the incentive constraint, the Lagrangian is given by

$$\mathcal{L}^w_t = (1+\lambda^w_t)\left(\mu^w_{kt}\frac{Q_t k^w_t}{n^w_t} + \mu^w_{bt}\frac{d^w_t}{n^w_t} + \nu^w_{bt}\right) - \lambda^w_t \theta \left[\frac{d^w_t}{n^w_t} + 1 + \omega Max\left(\frac{Q_t k^w_t}{n^w_t} - \frac{d^w_t}{n^w_t} - 1, 0\right)\right].$$

The first order conditions for $\frac{Q_t k_t^w}{n_t^w}$ and $\frac{d_t^w}{n_t^w}$ are

$$(1+\lambda_t^w)\mu_{kt}^w = \begin{cases} \lambda_t^w \theta\omega, \text{ if } Q_t k_t^w > d_t^w + n_t^w \\ 0, \text{ otherwise} \end{cases}$$

$$\begin{aligned}(1+\lambda_t^w)\mu_{bt}^w &\leq \begin{cases} \lambda_t^w \theta(1-\omega), \text{ if } Q_t k_t^w > d_t^w + n_t^w \\ \lambda_t^w \theta, \text{ otherwise} \end{cases} \\ \text{where } &= \text{ holds if } d_t^w > 0.\end{aligned}$$

We choose the parameters so that $0 < \mu_{kt}^w$ holds in the neighborhood of the steady state. Moreover we numerically show $0 < \mu_{kt}^w$ always holds in our dynamic equilibrium. Then we learn $Q_t k_t^w > d_t^w + n_t^w$ in order to be consistent with the wholesale bank's choice. Then

$$\mu_{kt}^w = \frac{\lambda_t^w}{1+\lambda_t^w}\theta\omega, \tag{46}$$

$$(1+\lambda_t^w)\mu_{bt}^w \leq \lambda_t^w \theta(1-\omega), \text{ where } = \text{ holds if } d_t^w > 0. \tag{47}$$

Thus we learn $0 < \mu_{kt}^w$ implies

$$\lambda_t^w > 0, \text{binding incentive constraint,}$$

and

$$\mu_{kt}^w < \theta\omega.$$

Also from $(46, 47)$ with $\lambda_t^w > 0$, we learn

$$\omega\mu_{bt}^w \leq (1-\omega)\mu_{kt}^w, \text{ where } = \text{ holds if } d_t^w > 0.$$

This implies (34) in the text.

The retail bank chooses $\left(\frac{(Q_t+f_t^r)k_t^r}{n_t^r}, \frac{d_t^r}{n_t^r}\right)$ to maximize Tobin's Q (17) subject to the incentive constraint (18). Because we learn wholesale banks borrow $Q_t k_t^w > d_t^w + n_t^w$ in our equilibrium, retail banks lend in the interbank market:

$$(Q_t + f_t^r)k_t^r < d_t^r + n_t^r.$$

Letting λ_t^r be Lagrangian multiplier of the incentive constraint, the Lagrangian is given by

$$\mathcal{L}_t^r = (1+\lambda_t^r)\left(\mu_{kt}^r \frac{(Q_t+f_t^r)k_t^r}{n_t^r} + \mu_{bt}^r \frac{d_t^r}{n_t^r} + \nu_{bt}^r\right) - \lambda_t^r \theta\left(\frac{d_t^r}{n_t^r} + 1\right).$$

The first order conditions for $\frac{(Q_t+f_t^r)k_t^r}{n_t^r}$ and $\frac{d_t^r}{n_t^r}$ are

$$(1+\lambda_t^r)\mu_{kt}^r \leq 0, \text{ where } = \text{holds if } k_t^r > 0,$$

$$\begin{aligned} (1+\lambda_t^r)\mu_{bt}^r &= \lambda_t^r\theta, \text{ or} \\ \mu_{bt}^r &= \frac{\lambda_t^r}{1+\lambda_t^r}\theta. \end{aligned}$$

We choose the parameters so that $0 < \mu_{bt}^r$ holds in the neighborhood of the steady state, and numerically show $0 < \mu_{bt}^r$ always holds in our dynamic equilibrium. Then we learn

$$\lambda_t^r > 0, \text{ binding incentive constraint,}$$

and

$$\mu_{bt}^r < \theta.$$

Then the argument in the text follows.

6.2 Appendix B: Steady State of the Economy without Run

To study the steady state it is convenient to introduce a new variable that measures the total return net worth of banks of type j as

$$x^j = \frac{1-\frac{W^j}{N^j}}{\sigma^j}, \tag{48}$$

where $j = w$ and r for wholesale and retail banks. Thus we learn

$$N^j = \frac{W^j}{1-\sigma^j x^j}, \text{ for } x^j \in \left(0, \frac{1}{\sigma^j}\right) \tag{49}$$

Then from $(35, 36)$, we get

$$x^j = (R_k^j - R_b)\frac{(Q+f^j)K^j}{N^j} + (R_b - R)\frac{D^j}{N^j} + R_b. \tag{50}$$

Let $\psi^j = V^j/n^j$ be Tobin's Q of banks of type j. Then from $(17, 19, 20, 21)$, we have

$$\begin{aligned} \psi^j &= \beta\left(1-\sigma+\sigma\psi^j\right)\left[(R_k^j - R_b)\frac{(Q+f^j)K^j}{N^j} + (R_b - R)\frac{D^j}{N^j} + R_b\right] \\ &= \frac{(1-\sigma^j)\,\beta x^j}{1-\beta\sigma^j x^j}. \end{aligned}$$

Then from $(25, 28)$ and (18) with equality, we have

$$\begin{aligned}
\frac{QK^w}{N^w} + \frac{1-\omega}{\omega}\frac{D^w}{N^w} &= \frac{1}{\theta\omega}\frac{(1-\sigma^w)\,\beta x^w}{1-\beta\sigma^w x^w} - \frac{1-\omega}{\omega} = \phi^w(x^w) && (51)\\
\frac{(Q+f^r)K^r}{N^r} + \frac{-B^r}{N^r} &= \frac{1}{\theta}\frac{(1-\sigma^r)\,\beta x^r}{1-\beta\sigma^r x^r} = \phi^r(x^r). && (52)
\end{aligned}$$

From (7), we learn

$$\beta R = 1. \tag{53}$$

Then from $(27, 34, 50)$, we find R^b, R^w and Q

$$\begin{aligned}
R_b &= R + \frac{1}{\phi^r(x^r)}\left(x^r - R\right) && (54)\\
R_k^w &= R_b + \frac{1}{\phi^w}\left(x^w - R_b\right) && (55)\\
Q &= \frac{Z}{R_k^w - 1} && (56)
\end{aligned}$$

Then from the conditions of retail banks and households, we can derive their demand for capital K^r and K^h as

$$\frac{Z+Q}{Q+\alpha^r K^r} = R_b \tag{57}$$

$$\beta\frac{Z+Q}{Q+\alpha^h K^h} = 1. \tag{58}$$

The interbank loan of retail banks B is

$$B = \phi^r N^r\left(x^r\right) - \left(Q + \alpha^r K^r\right)K^r. \tag{59}$$

The market clearing condition for capital is given as

$$K^h + K^w + K^r = \overline{K}. \tag{60}$$

The household consumption can be found from the goods market clearing condition as

$$\begin{aligned}
& Z_t\left(1+W^h\right) + W^w + W^r && (61)\\
= \; & C^h + (1-\sigma^w)\,x^w N^w + (1-\sigma^r)\,x^r N^r + \frac{\alpha^h}{2}\left(K^h\right)^2 + \frac{\alpha^r}{2}\left(K^r\right)^2.
\end{aligned}$$

For the case in which wholesale banks only raise external funds in interbank market, we have

$$\phi^w N^w (x^w) = QK^w = Q(\overline{K} - K^h - K^r). \tag{62}$$

For the market clearing condition for the interbank market, we get

$$\begin{aligned} (\phi^w - 1) N^w (x^w) &= B \\ &= \phi^r N^r (x^r) - (Q + \alpha^r K^r) K^r \end{aligned} \tag{63}$$

Then we can find (x^w, x^r) to satisfy these market clearing conditions of capital and interbank loans. We have to check that the condition for $D^w = 0$ in (34) is satisfied in the equilibrium as

$$\omega \left[R_b(x^r) - \frac{1}{\beta} \right] < (1 - \omega) \left[R_k^w(x^w, x^r) - R_b(x^r) \right].$$

For the case in which wholesale banks raise external funds in both retail deposit and interbank markets, we have

$$\omega \left[R_b(x^r) - \frac{1}{\beta} \right] = (1 - \omega) \left[R_k^w(x^w, x^r) - R_b(x^r) \right]. \tag{64}$$

Using this together with (51) into (55) we get R_k^w as a function of x^w only

$$R_k^w (x^w) - R = \theta \frac{(x^w - R) (R - \sigma^w x^w)}{(1 - \sigma^w) x^w}$$

Hence equation (64) can be used to solve for values x^r associated with x^w that satisfy the market clearing condition in the interbank loans

$$\frac{(x^r - R) (R - \sigma^r x^r)}{(1 - \sigma^r) x^r} = (1 - \omega) \frac{(x^w - R) (R - \sigma^w x^w)}{(1 - \sigma^w) x^w} \tag{65}$$

notice that a little algebra shows that for values of x^w that make the right hand side positive, *i.e.* $R < x^w < \frac{R}{\sigma^w}$, the two roots of this equation, when they are real, both satisfy $R < x^r < \frac{R}{\sigma^r}$. So as long as they imply positive N^r they are both possible.

Once we find candidate values for x^r we need to check market clearing in the capital market. Since wholesale are not fully specialized we need to get their capital demand from

$$\phi^w N^w = QK^w + \frac{1 - \omega}{\omega} (QK^w - B - N^w).$$

Together with $(59, 60)$, we get

$$\begin{aligned}
QK^w &= Q(\overline{K} - K^h - K^r) && (66)\\
&= (\omega\phi^w + 1 - \omega)\, N^w\,(x^w) + (1-\omega)B\\
&= [\omega\phi^w + 1 - \omega]\, N^w\,(x^w) + (1-\omega)\left\{\phi^r N^r\,(x^r) - (Q + \alpha^r K^r)\, K^r\right\}.
\end{aligned}$$

Then, we can find (x^w, x^r) to satisfy these two market clearing conditions $(65, 66)$. We have to check that the condition for $D^w > 0$ is satisfied in the equilibrium as

$$\begin{aligned}
0 &< D^w = QK^w - N^w\,(x^w) - B, \text{ or}\\
0 &< [\phi^w - 1]\, N^w\,(x^w) - [\phi^r(w^r) N^r\,(x^r) - (Q + \alpha^r K^r)\, K^r].
\end{aligned}$$

6.3 Effect of varying ω on the steady state

Let's study equation (51)

$$\begin{aligned}
\phi^w\,(w^w) &= \frac{1}{\theta\omega}\left[\frac{1-\sigma^w}{\sigma^w}\frac{\beta(1-w^w)}{1-\beta(1-w^w)} - \theta(1-\omega)\right]\\
&= 1 + \frac{1}{\theta\omega}\left[\frac{1-\sigma^w}{\sigma^w}\frac{\beta(1-w^w)}{1-\beta(1-w^w)} - \theta\right].
\end{aligned}$$

Then, because $\phi^w\,(w^w) > 1$, we see that keeping w^w constant, that is neglecting the general equilibrium effect on net worth, the leverage multiple is a decreasing function of ω :

$$\frac{\partial\phi^w\,(w^w)}{\partial\omega} < 0.$$

At the same time, lower levels of w^w, i.e. higher levels of steady state net worth, increases leverage as

$$\frac{\partial\phi^w\,(\omega, w^w)}{\partial w^w} < 0.$$

Notice that the general equilibrium adjustment of net worth following a change in ω works entirely through the shift in the demand for interbank borrowing and capital that is generated by the direct effect of ω on ϕ^w : as ϕ^w increases asset prices Q increases and interbank borrowing costs R^b decrease hence the effects on net worth are not straightforward and could partially offset the initial increase in ϕ^w if the increase in prices strong enough.

References

[1] Adrian, T., Colla, P., and Shin, H., 2012. Which Financial Frictions? Paring the Evidence from Financial Crisis of 2007-9. Forthcoming in NBER Macroeconomic Annual.

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

[3] Allen, F., and Gale, D., 2007. Understanding Financial Crises. Oxford University Press.

[4] Angeloni, I., and E. Faia, 2013, Capital Regulation and Monetary Policy with Fragile Banks, Journal of Monetary Policy 60, 3111-382.

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

[6] Bernanke, B., and Gertler, M., 1989. Agency Costs, Net Worth and Business Fluctuations. American Economic Review 79, 14-31.

[7] Bianchi, J., 2011. Overborrowing and Systemic Externalities in the Business Cycle. American Economic Review 101, 3400-3426.

[8] Boissay F., F. Collard, and F. Smets, 2013, Booms and Systemic Banking Crises, mimeo.

[9] Brunnermeier, M. K., and M. Oemke. 2013 Maturity Rat Race. Journal of Finance, 68: 483-521.

[10] Brunnermeier, M. K., and Pedersen, L., 2009. Market Liquidity and Funding Liquidity. Review of Financial Studies 22, 2201-2238.

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

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

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

[14] Covitz, D., N. Liang and G. Suarez. 2013. Evolution of a Financial Crisis: Collapse of the Asset-Backed Commercial Paper Market. Journal of Finance, 68: 815-

[15] Curdia, V., and M. Woodford. 2010. Credit Spreads and Monetary Policy. Journal of Money, Credit and Banking, 42(6), 3-35.

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

[17] Eggertsson, G., and P. Krugman. 2012. Debt, Deleveraging, and Liquidity Trap. Quarterly Journal of Economics, 1469-1513.

[18] Ennis, H. and T. Keister, 2003, Economic Growth, Liquidity and Bank Runs, Journal of Economic Theory 109, 220-245.

[19] Farmer, R., 1999. The Macroeconomics of Self-Fulfilling Prophecies. MIT Press.

[20] Ferrante, F. 2015. A Model of Endogenous Loan Quality and the Collapse of the Shadow Banking System. Finance and Economics Discussion Series 2015-021, Federal Reserve Board.

[21] Friedman, M., and Schwartz, A., 1963. A Monetary History of the United States, 1867-1960. Princeton University Press.

[22] Gertler, M., and Karadi, P., 2011. A Model of Unconventional Monetary Policy, Journal of Monetary Economics, January.

[23] Gertler, M., and Kiyotaki, N., 2011. Financial Intermediation and Credit Policy in Business Cycle Analysis. In Friedman, B., and Woodford, M. (Eds.), Handbook of Monetary Economics. Elsevier, Amsterdam, Netherlands.

[24] Gertler, M., Kiyotaki, N., and Queralto, A., 2012. Financial Crises, Bank Risk Exposure and Government Financial Policy, Journal of Monetary Economics 59, S17-S34.

[25] Gertler, M., N. Kiyotaki. 2015. Banking, Liquidity and Bank Runs in an Infinite Horizon Economy, American Economic Review, forthcoming.

[26] Gilchrist, S., and E. Zakrajsek. 2012. Credit Spread and Business Cycle Fluctuations, American Economic Review.

[27] Goldstein, I., and A. Pauzner. 2005. Demand-Deposit Contracts and the Probability of Bank Runs. Journal of Finance 60, 1293-1327.

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

[29] Guerrieri, V., and G. Lorenzoni. 2011. Credit Crises, Precautionary Savings and the Liquidity Trap. NBER Working Paper 17583.

[30] Gurley, J., and E. Shaw. 1960. Money in Theory of Finance. Washington, D.C.: Brookings Institution.

[31] He, Z. and A. Krishnamurthy, 2013. Intermediary Asset Pricing, American Economic Review.

[32] He, Z. and A. Krishnamurthy, 2014. A Macroeconomic Framework for Quantifying Systemic Risk. Working Paper, University of Chicago and Stanford University.

[33] Holmstrom, B. and J. Tirole, 2011. Inside and Outside Liquidity. MIT Press.

[34] Kacperczyk, M., and P. Schnabl. 2010. When Safe Proved Risky: Commercial Paper during the Financial Crisis of 2007-2009. Journal of Economic Perspectives, 24(1), 29-50.

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

[36] Krishnamurthy, A., S. Nagel and D. Orlov. 2014. Seizing Up Repo. Journal of Finance. Forthcoming.

[37] Lorenzoni, G., 2008. Inefficient Credit Boom. Review of Economic Studies 75, 809-833.

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

[39] Martin, A., D. Skeie and E.L. Von Thadden. Fragility of Short-Term Secured Funding. Journal of Economic Theory, 149: 15-42.

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

[41] Morris, S. and H.S. Shin. 1998. Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks. American Economic Review 88, 587-97.

[42] Phillipon, T. 2013. Has the U.S. Finance Industry Become Less Efficient?

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

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

Table 1: CALIBRATION

	Households	
β	discount rate	.99
α^h	Intermediation cost	.03
W^h	Endowment	.006
	Retail Banks	
σ^r	Survival Probability	.95
α^r	Intermediation cost	.0075
W^r	Endowment	.0008
θ	Divertable proportion of assets	.25
	Wholesale Banks	
σ^w	Survival Probability	.9
α^w	Intermediation cost	0
W^w	Endowment	.0004
ω	Shrinkage of divertable proportion of assets	.5
	Production	
σ_z	std of dividends	.05
ρ_z	autocorrelation of dividends	.9

Table 2: STEADY STATE

	Steady State	
Q	price of capital	1
K^r	retail intermediation	.4
K^w	wholesale intermediation	.4
R^b	Annual interbank rate	1.052
R	Annual deposit rate	1.04
R_w^k	Annual wholesale return on capital	1.064
ϕ^w	wholesale leverage	25
ϕ^r	retail leverage	10
Y	output	.0225
C^h	consumption	.0168
N^r	retail banks networth	.0785
N^w	wholesale banks networth	.0160

Figure 1: Modes of Financial Intermediation

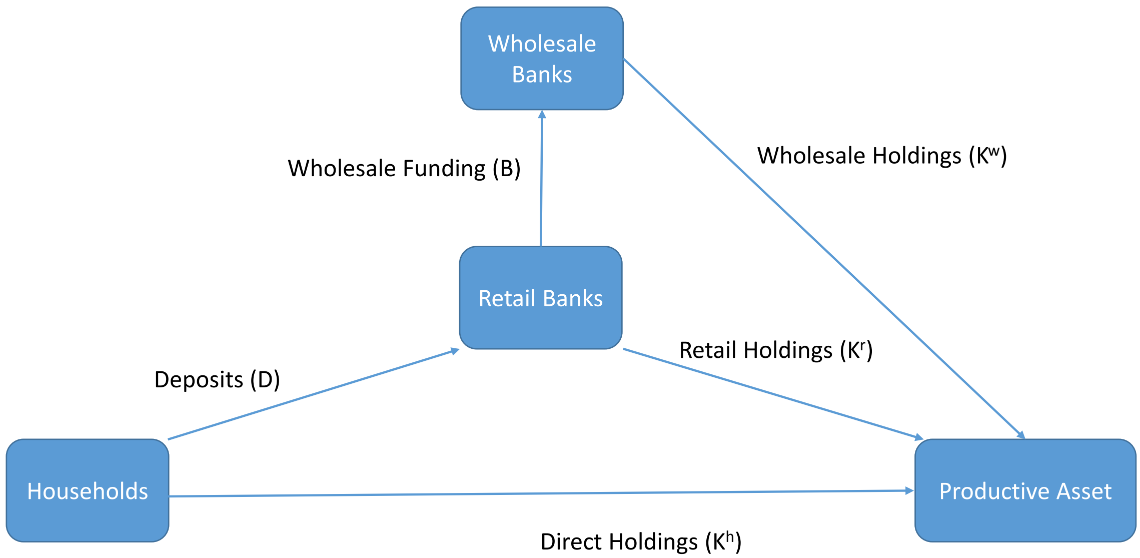


Figure 2: Wholesale Intermediation

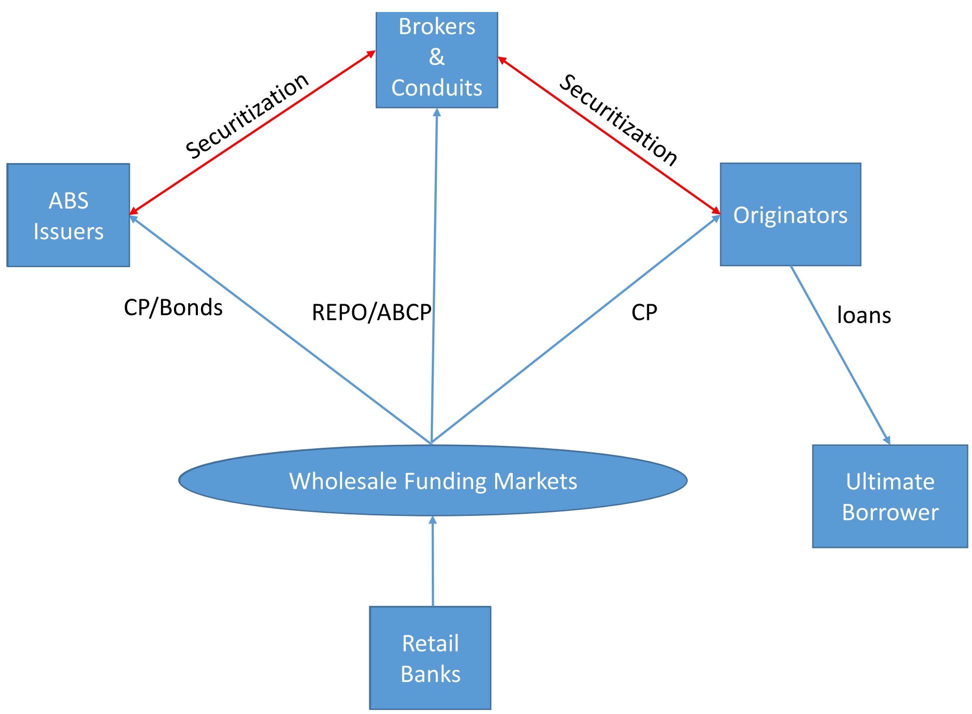
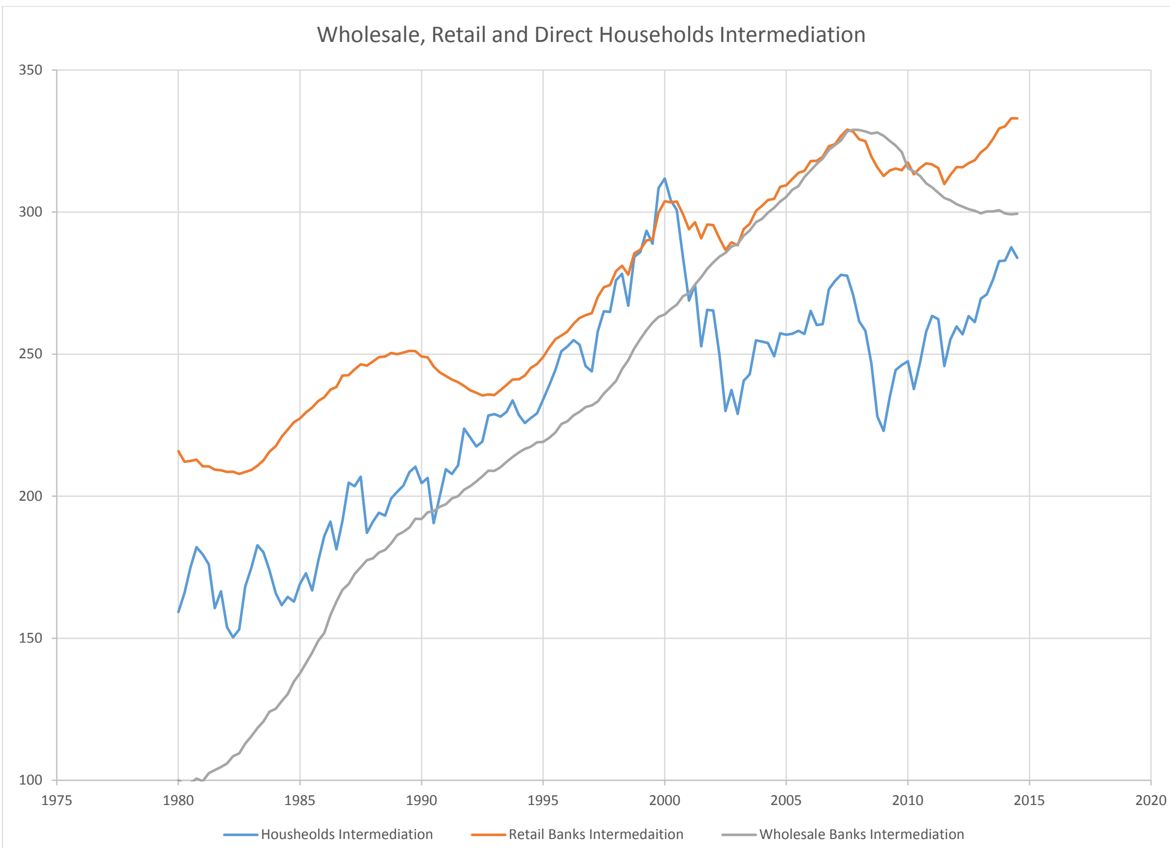
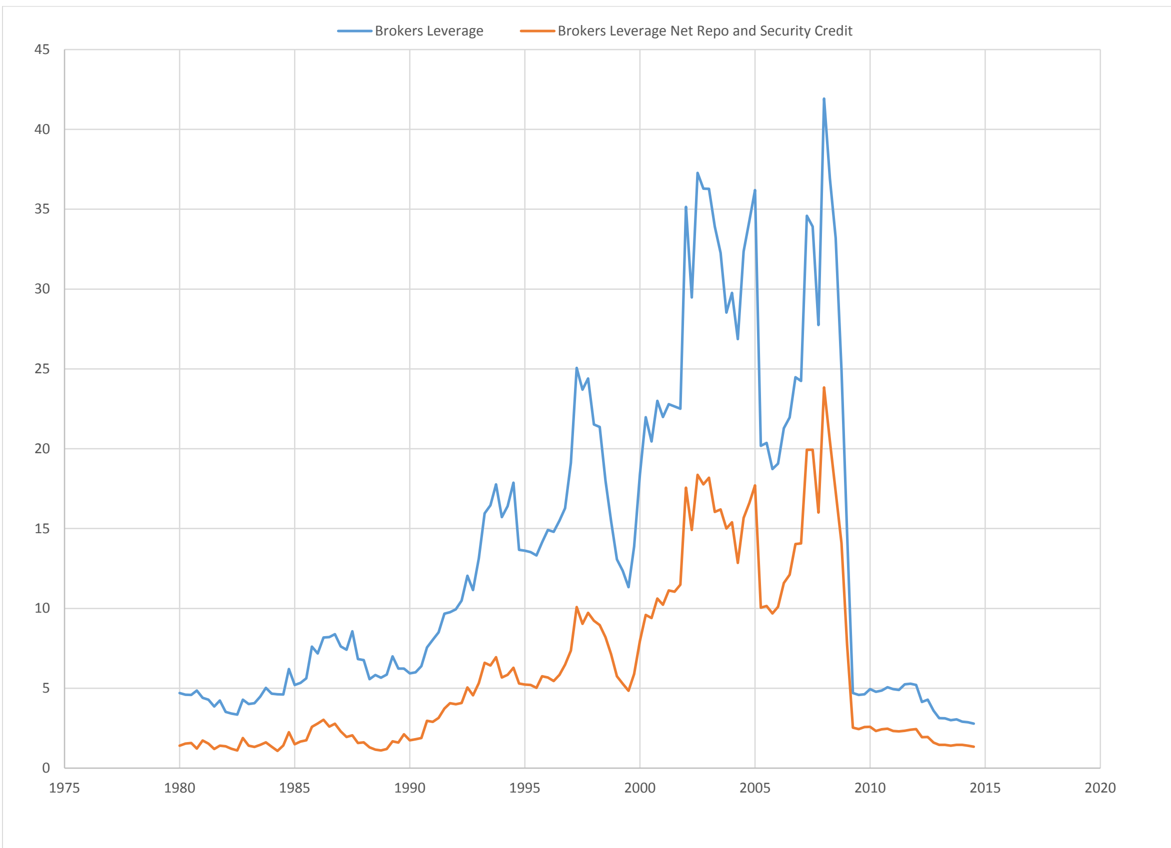


Figure 3: Intermediation by Sector



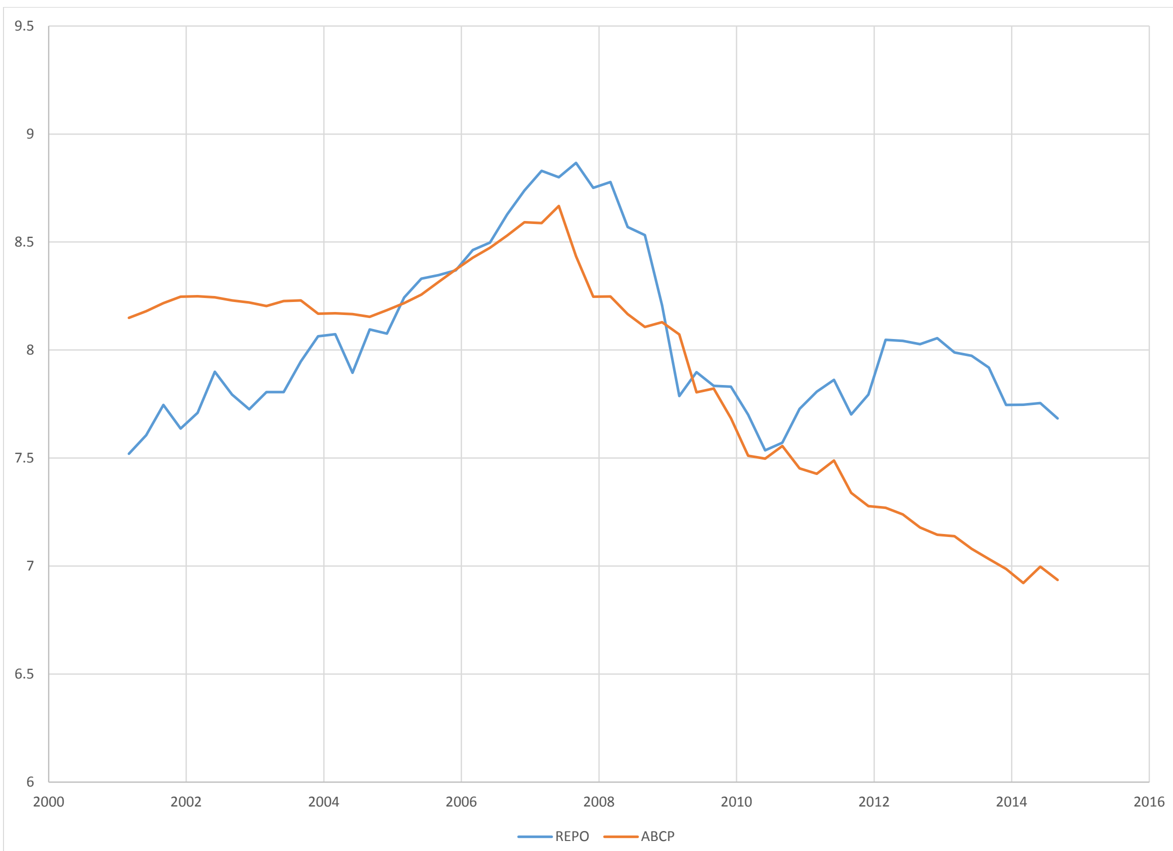
The graph shows the evolution of credit intermediated by the three different sectors. Nominal data from the flow of funds are deflated using the CPI and normalized so that the log of the normalized value of real wholesale intermediation in 1980 is equal to 1. The resulting time series are then multiplied by 100

Figure 4: Brokers Leverage



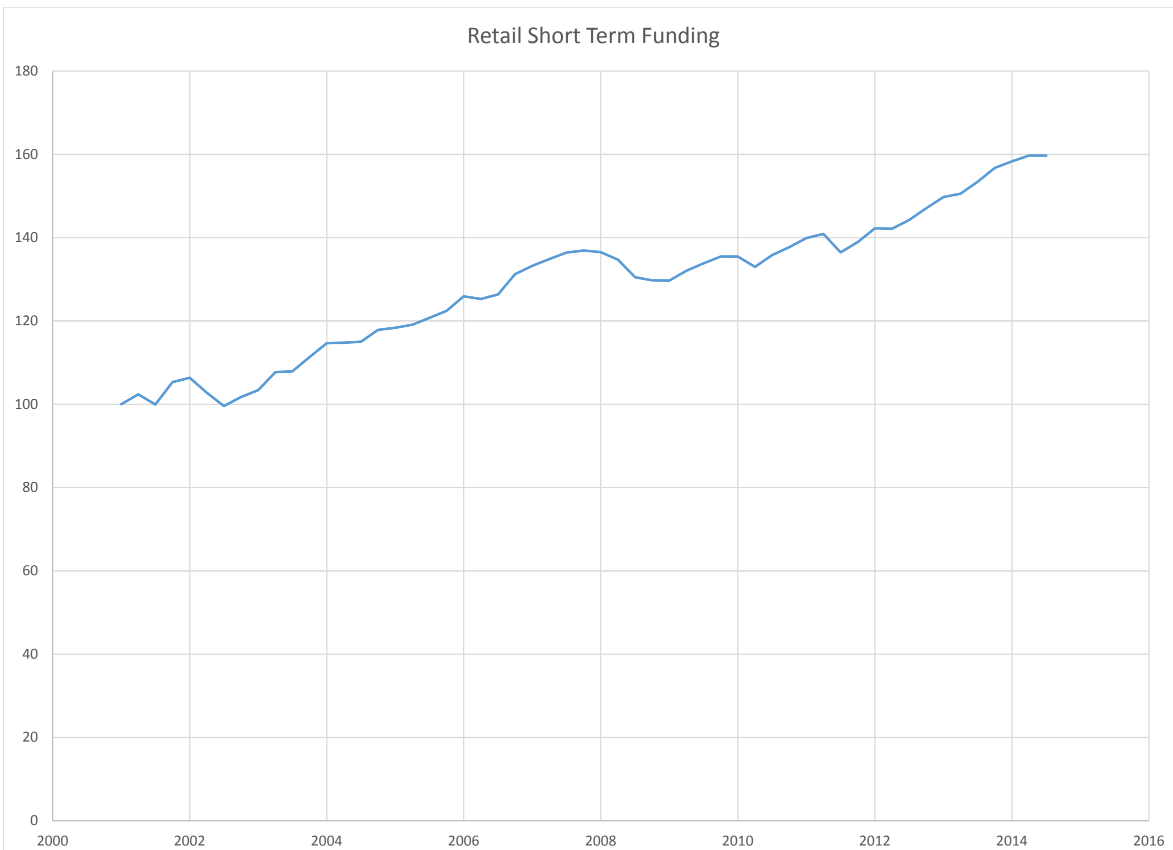
Leverage is given by the ratio of total assets over equity. Equity is computed from the flow of funds by subtracting liabilities other than "holding comanies equity investment" from total assets. The net position leverage computes assets by netting out long and short positions in REPO and Security Credit

Figure 5: Short Term Wholesale Funding



The graph shows the logarithm of the real value outstanding. Nominal values from Flow of FUnds are deflated using the CPI

Figure 6: Retail short term Funding



The graph shows the logarithm of the real value outstanding. Nominal values from Flow of Fnds are deflated using the CPI and normalized so that the log of the normalized value of retail short term funding in 2001 is equal to 100

Figure 7: Spreads

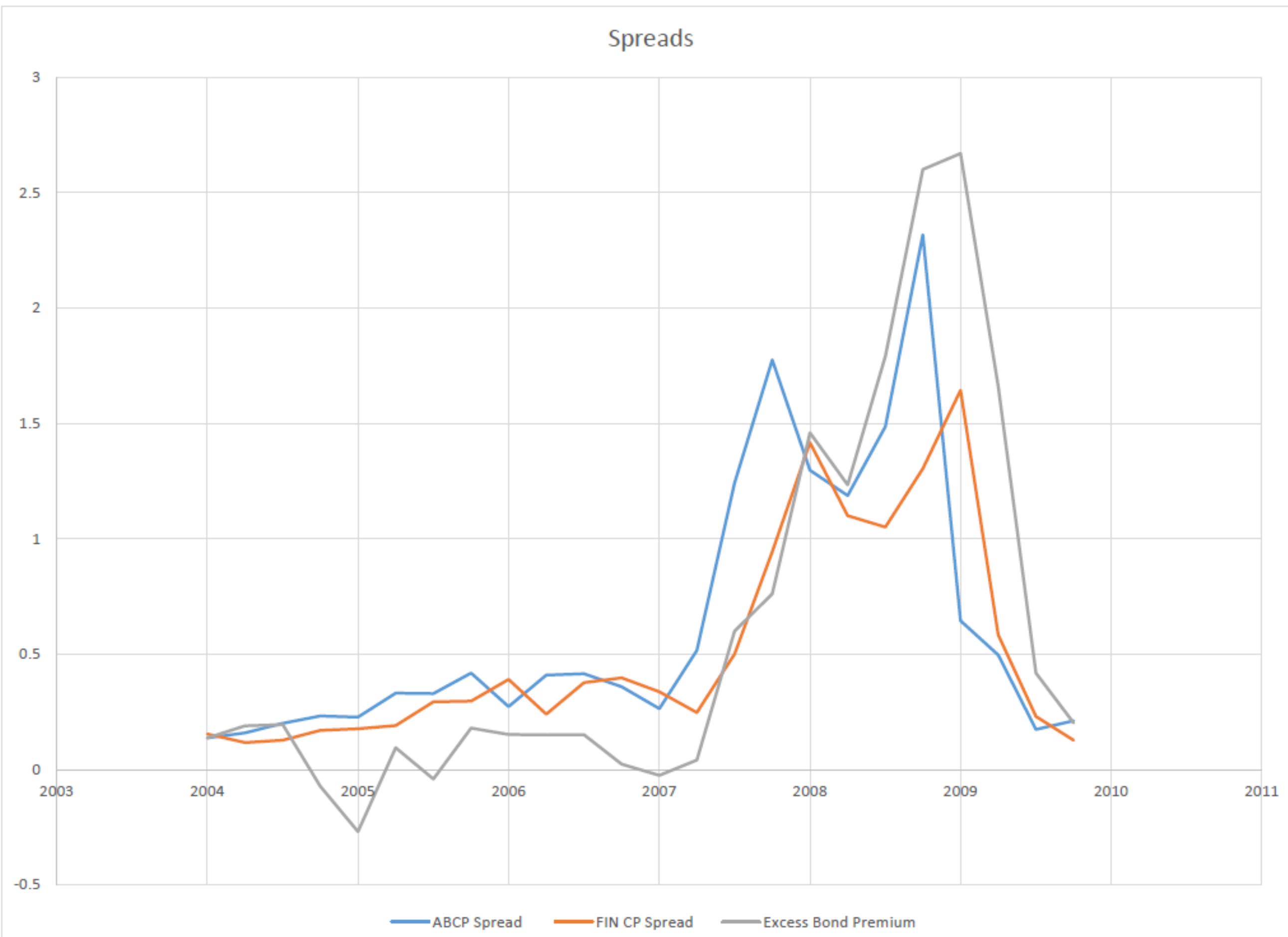


Figure 8: Investment Collapse

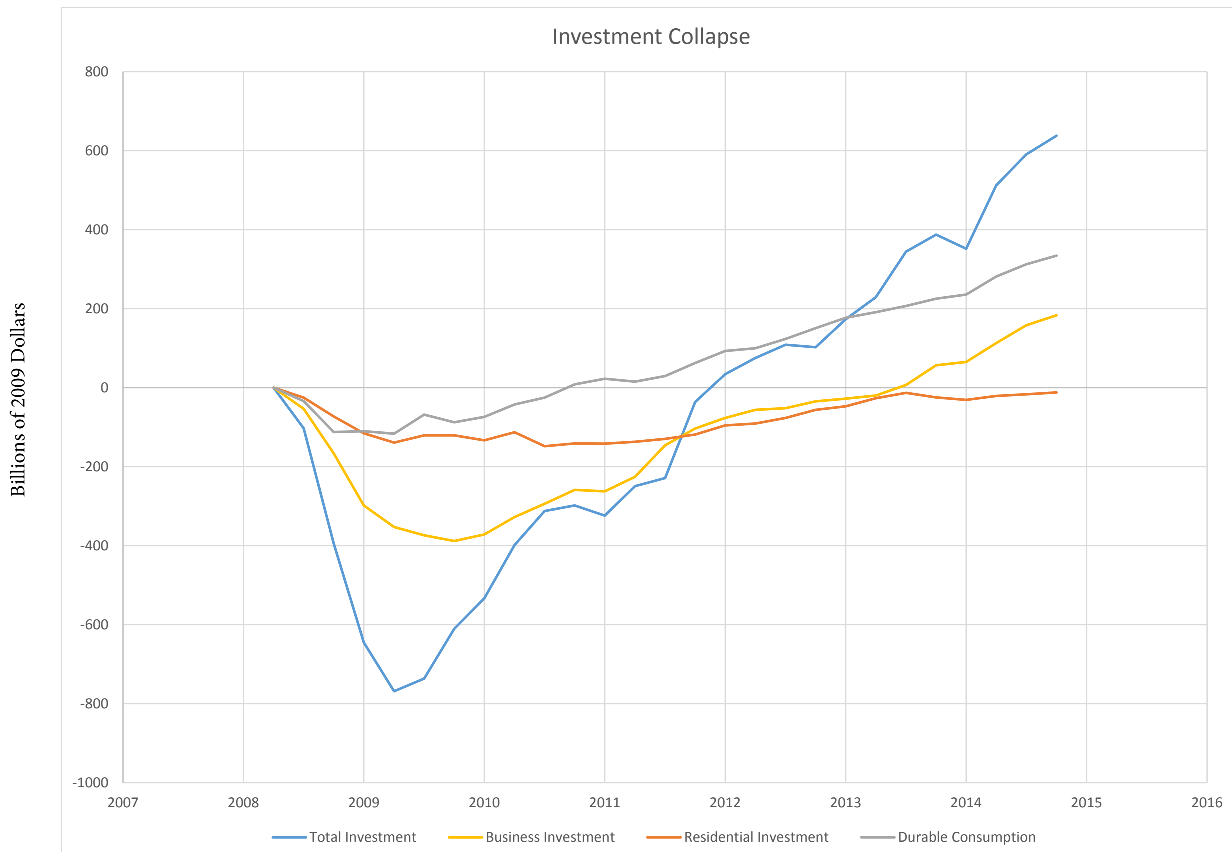


Figure 9: Comparative Statics: a reduction in ω

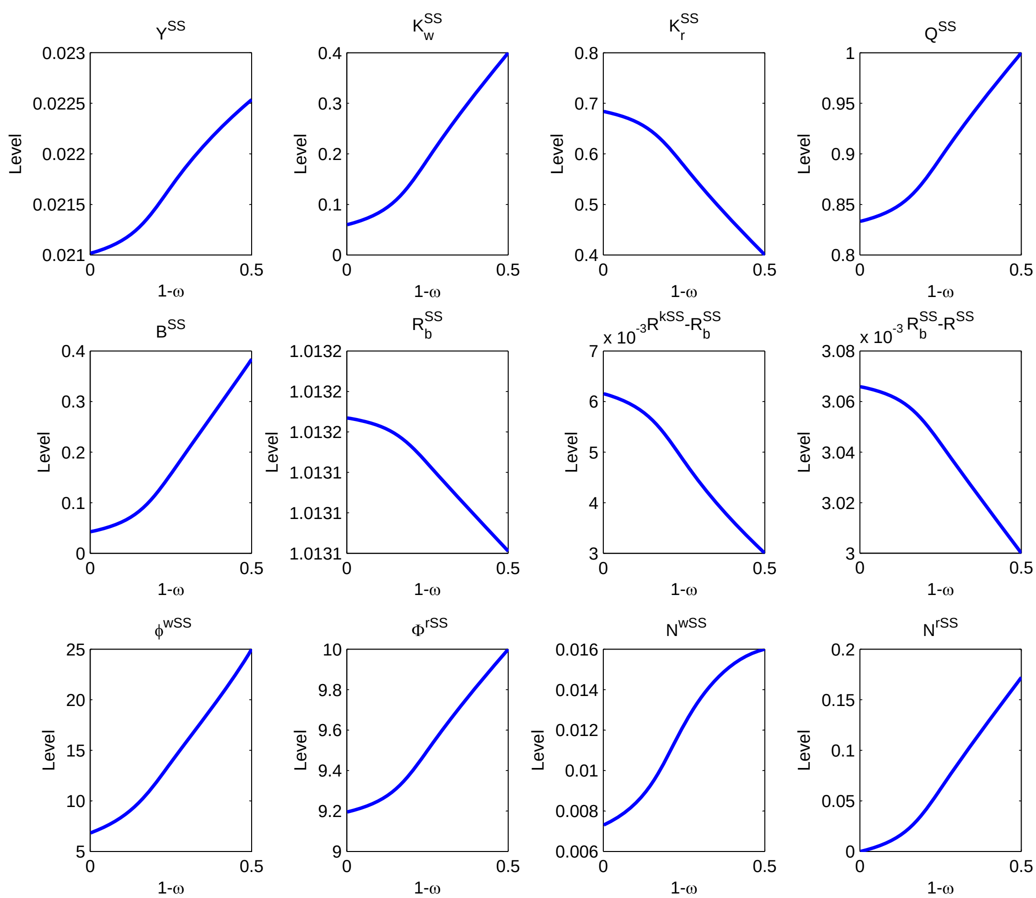


Figure 10: Low Frequency Dynamics in Financial Intermediation

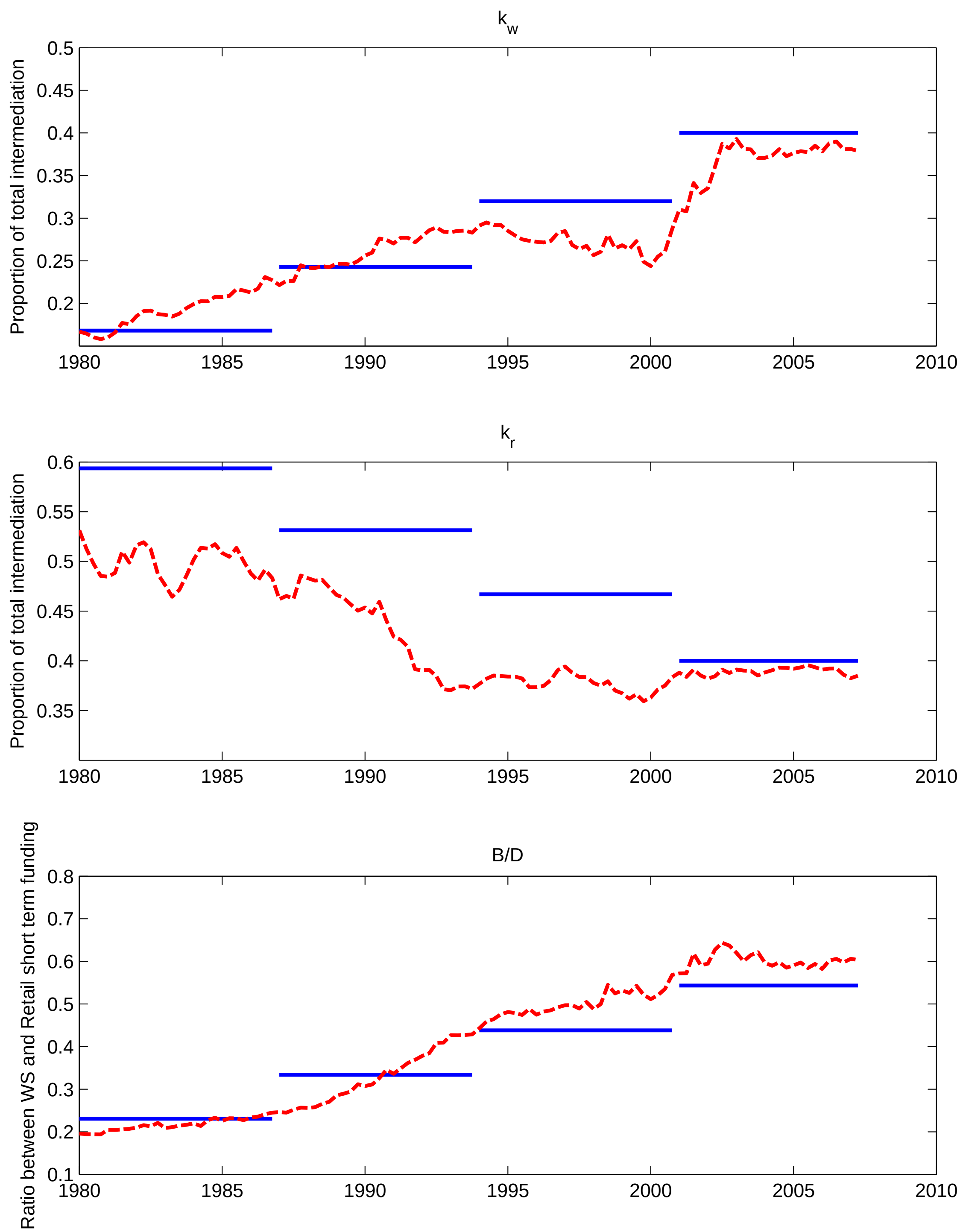


Figure 11: Low Frequency Growth in Leverage

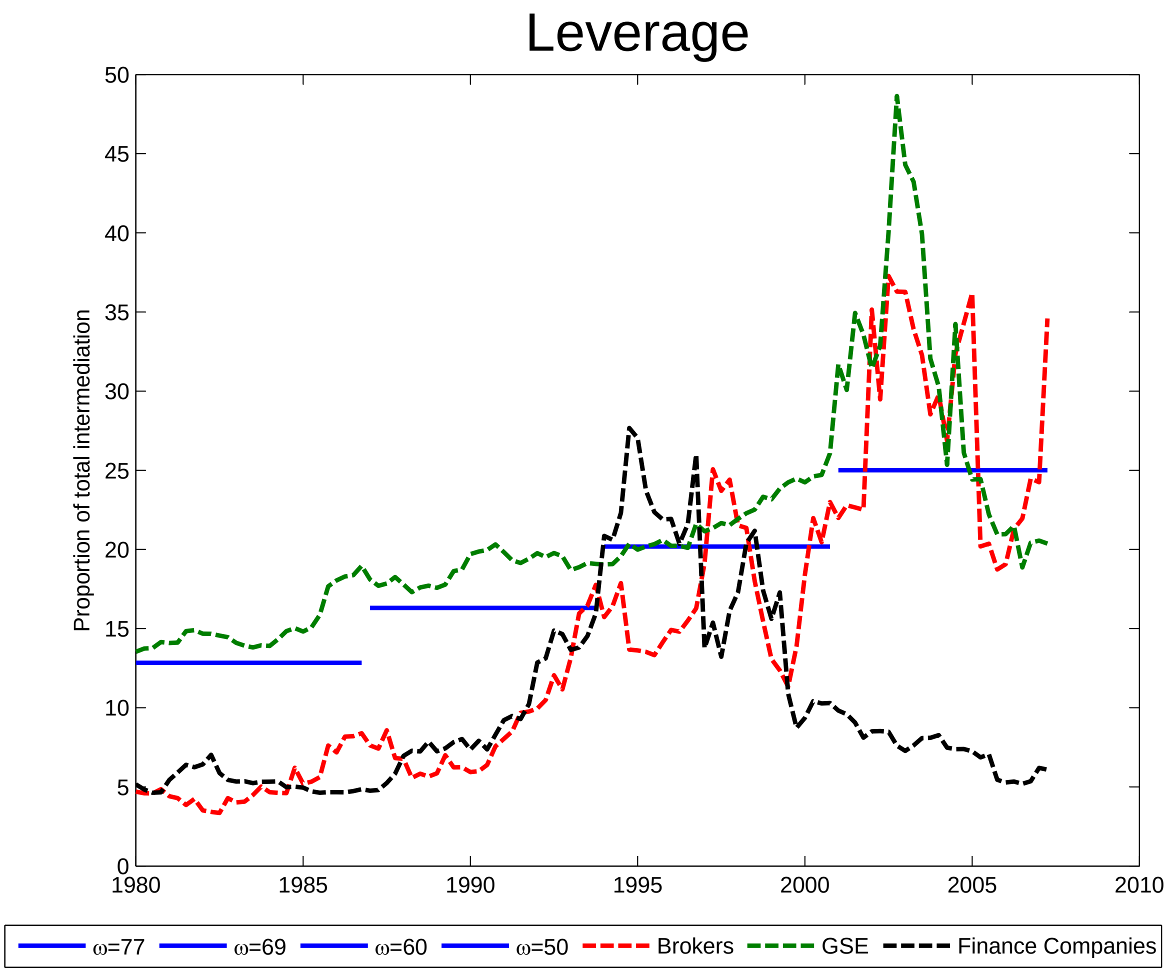


Figure 12: A recession before and after financial innovation (NO RUN EQUILIBRIUM)

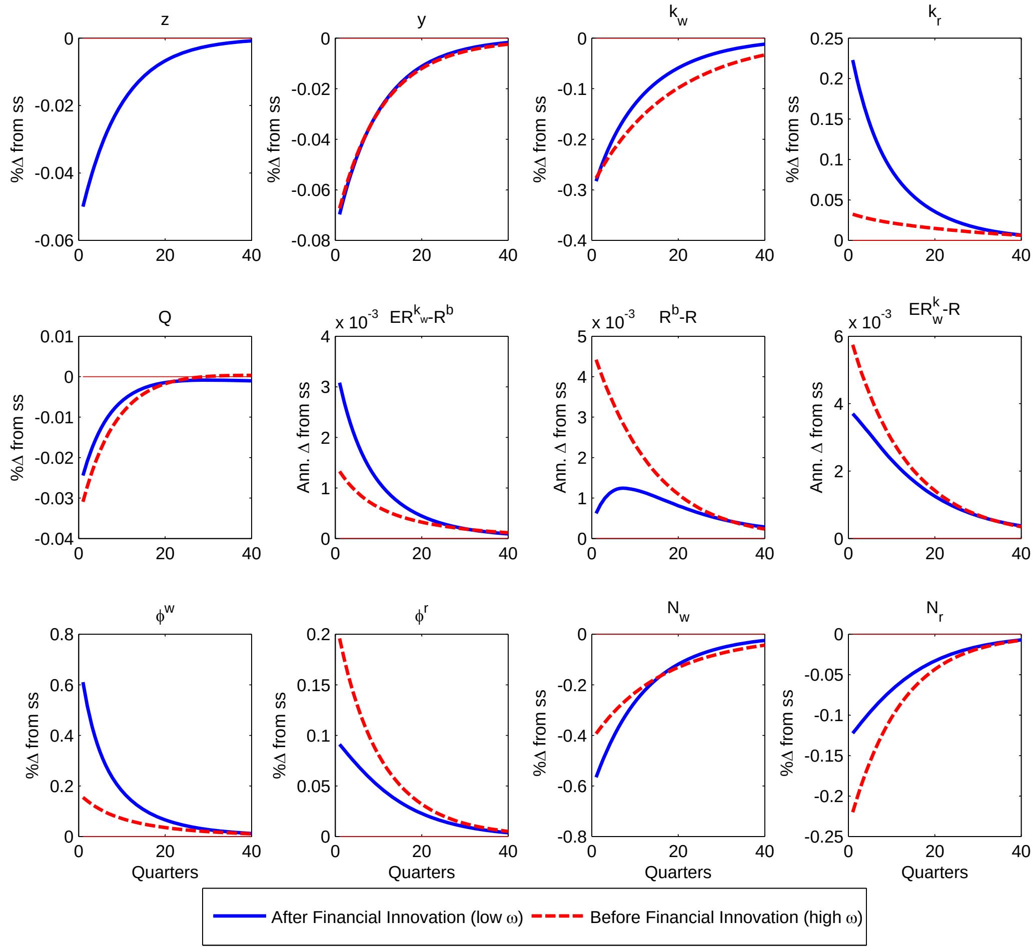


Figure 13: A Recession followed by a run on the entire banking system

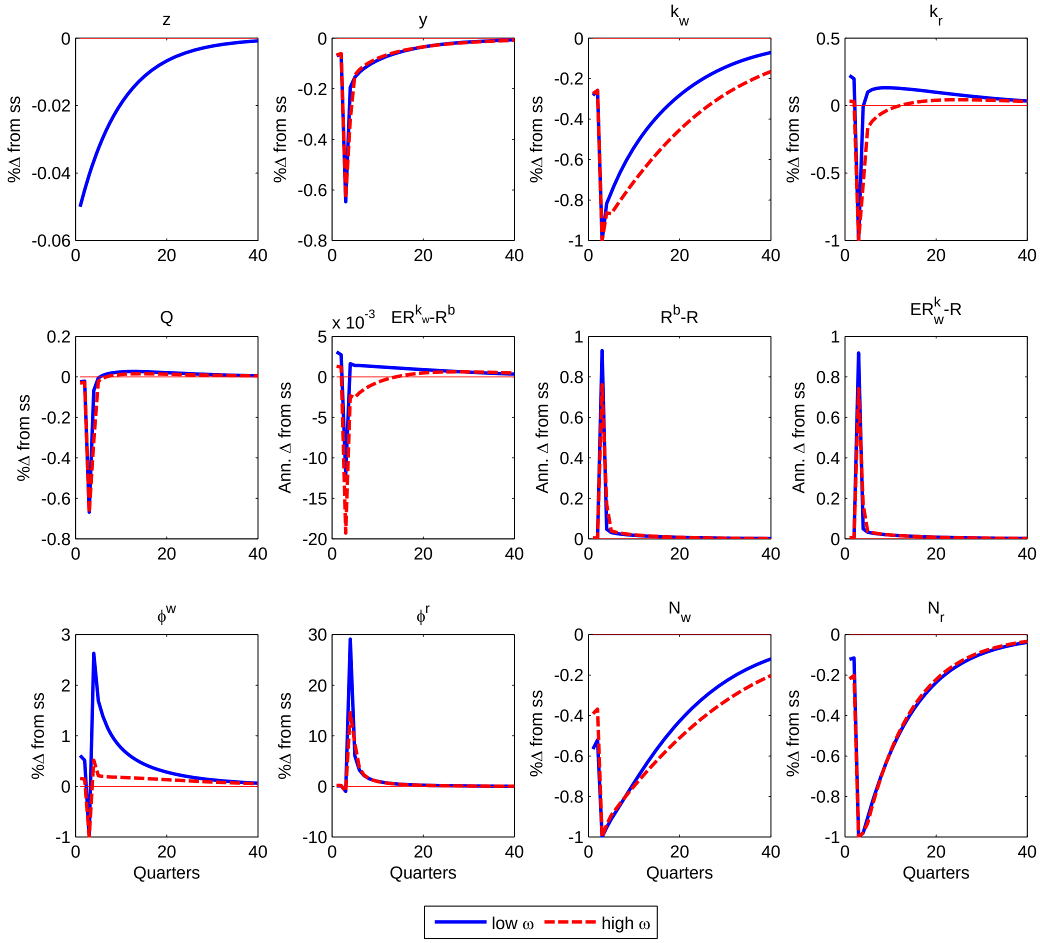


Figure 14: A recession followed by a run on wholesale bankers only

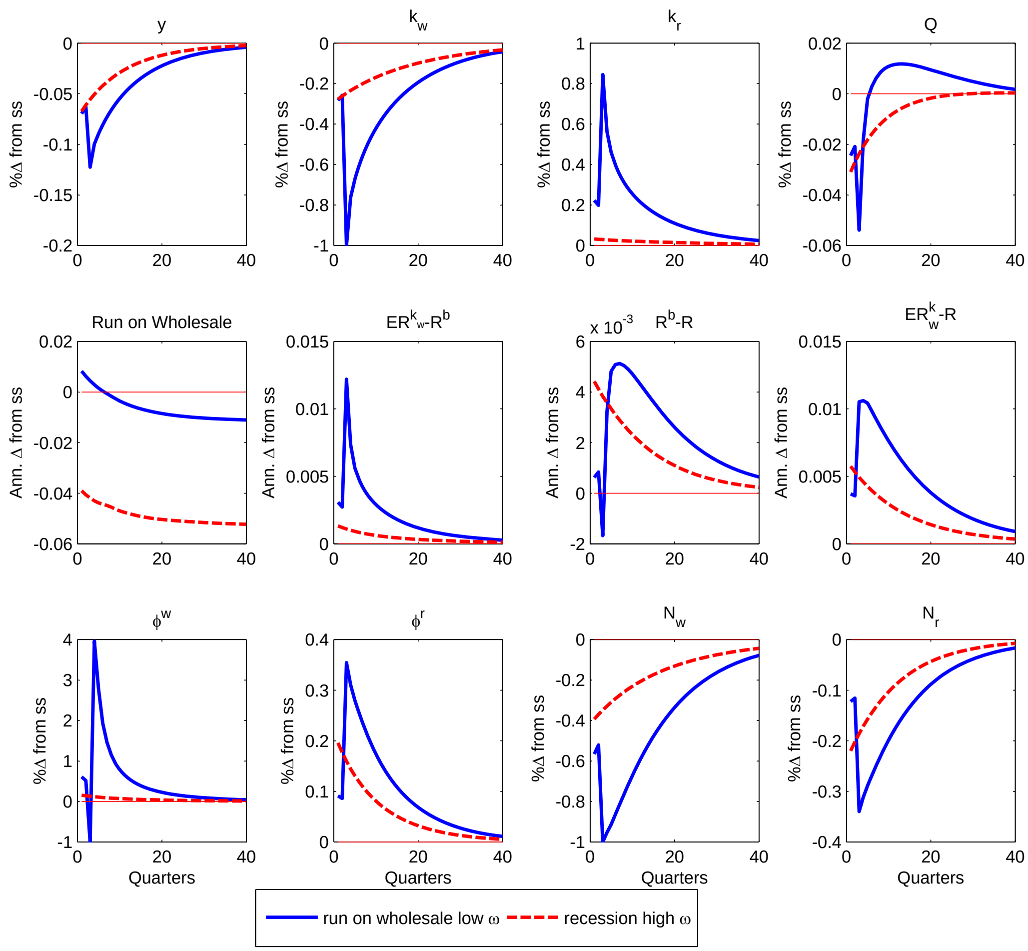


Figure 15: A recession in the model with anticipated runs

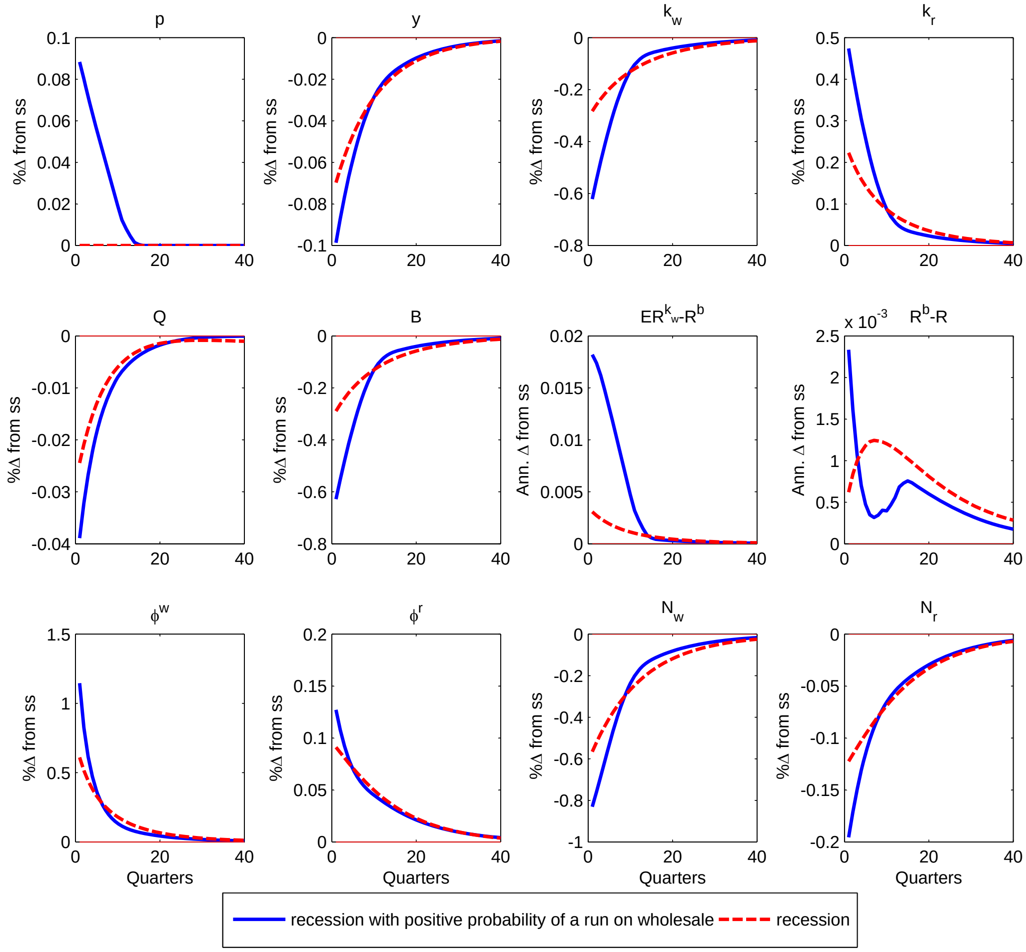


Figure 16: A recession followed by a run in the model with anticipated runs

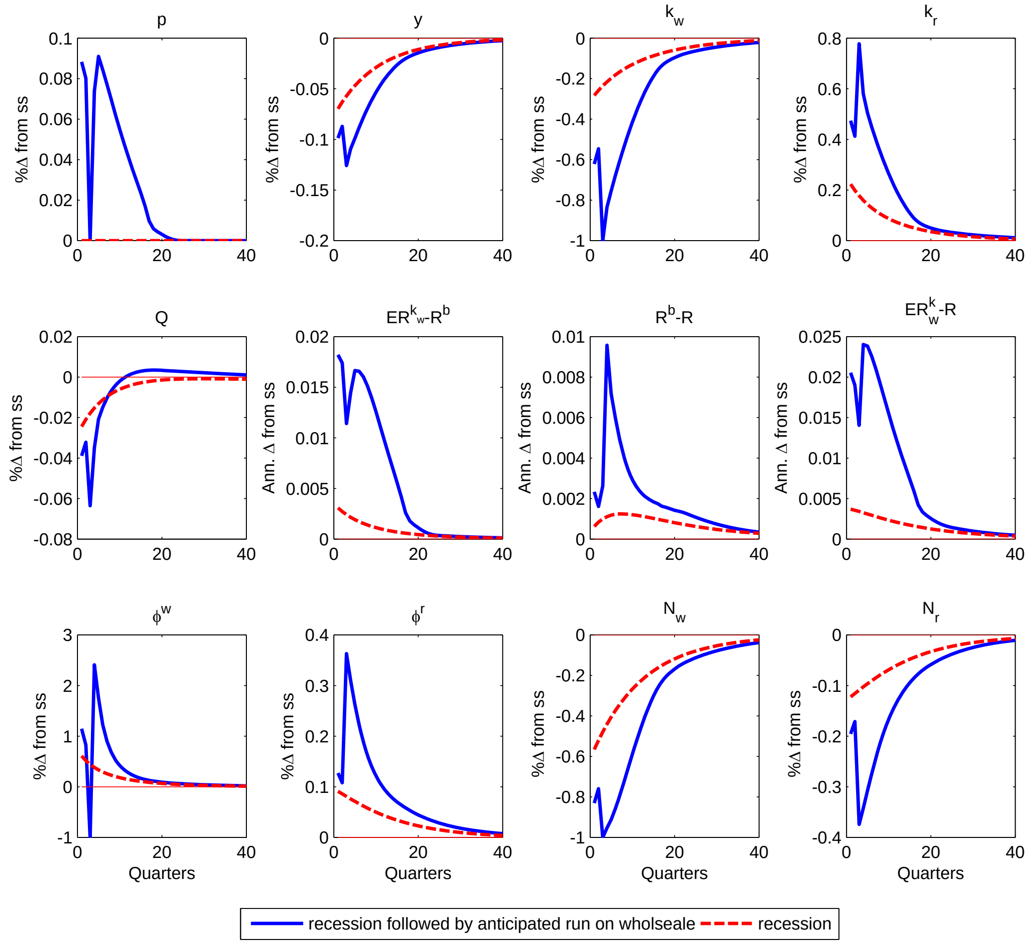


Figure 17: A recession followed by a run in the model with anticipated runs