Financial Disintermediation and Financial Fragility

Kosuke Aoki
The University of Tokyo

Kalin Nikolov
European Central Bank

November 2015

CARF Working Papers can be downloaded without charge from:
http://www.carf.e.u-tokyo.ac.jp/workingpaper/index.html

Working Papers are a series of manuscripts in their draft form. They are not intended for circulation or distribution except as indicated by the author. For that reason Working Papers may not be reproduced or distributed without the written consent of the author.
Financial Disintermediation and Financial Fragility*

Kosuke Aoki† Kalin Nikolov‡
University of Tokyo European Central Bank

November 2015

Abstract

This paper investigates how expanding the corporate bond market and the shadow banking sector affect the susceptibility of the financial system to crisis. We show that the corporate bond market can increase banking fragility although it also diminishes the impact of banking crises. Shadow banking allows higher financial system leverage and thus increases bank risk taking and fragility even further. Because it relies on bank capital for its operations, the shadow banking sector provides no funding diversification and cannot offset the real economy impact of a banking crisis.

JEL Classification: E44, G01, G23.

Key Words: Credit frictions, Shadow banking, Financial Crises, Asset Price Bubbles.

*We would like to thank our discussants Harald Uhlig, Martin Bodenstein, Michal Kowalik and Chia-Ying Chang, as well as participants at the Bundesbank-SAFE conference (Frankfurt, 2013), the 24th NBER East Asian Seminar on Economics Meetings (Auckland, 2013), the 4th Joint French Macro Workshop (Toulouse, 2013), the IBEFA Conference (Seattle, 2013) and the European Economic Association conference (Gothenburg, 2013) for excellent comments which helped us to improve the paper. We would also like to thank Emanuel Farhi, Jordi Gali, Hyun Shin and Philippe Weil for useful comments and discussions. Aoki thanks financial support from the Center for Advanced Research in Finance (CARF, University of Tokyo) and the JSPS Grant-in-Aid for Scientific Research No. 25380223.

†Department of Economics, University of Tokyo, Japan. Email: kaoki@e.u-tokyo.ac.jp
‡Research Department, European Central Bank, Frankfurt-am-Main, Germany. Email: kalin.nikolov@ecb.int
1 Introduction

Non-bank credit sources for firms and households have grown rapidly in the last two decades. Before the crisis, this process of disintermediation was seen as a benign or even beneficial phenomenon. Greenspan (1999) called the market for mortgage securitizations a useful ‘spare tire’ which insulated the US economy from the worst effects of the housing bust in the early 1990s. Having a diverse set of funding sources helped to maintain aggregate credit supply in the face of a contraction in bank lending.

The recent financial crisis changed researchers’ and policy-makers’ views of securitization and the repo market. These products and markets are seen by many as major culprits behind the leverage boom-bust cycle of the 2000s. In this paper we investigate which view is right. Does the growth of non-bank finance sources increase or decrease financial instability? Are they a ‘spare tire’ which contributes to the stability of aggregate credit supply or a source of systemic risk?

To answer the questions of the paper, we modify the limited commitment model with financial intermediation we developed in our earlier research (Aoki and Nikolov (2015)) by modelling two important financial innovations of the last 20 years - the shadow banking system and the corporate bond market. The first allows banks to expand leverage by shifting assets off-balance sheet. The second allows some loans to be conducted directly between borrowers and savers without the involvement of banks.

Both innovations reduce the aggregate shortage of bank capital in the financial intermediation process. This has the positive effect of increasing output and lending in the long run. However, the reduction in the scarcity of bank capital lowers lending-deposit spreads and reduces banks’ ‘franchise values’, making them more willing to hold riskier assets and increasing the likelihood of a banking crisis. In

\footnote{In line with the findings of the literature on bank competition and risk taking started by Keeley (1990), our model implies that the high profits from intermediation create ‘bank franchise value’. This prevents financial institutions from undertaking risky investment strategies.}
line with recent historical experiences, we find that the increase in financial fragility is especially pronounced when credit expands due to the increase in the size of the shadow banking sector.

Three main factors account for the greater severity of crises in economies with a large shadow banking sector. Firstly, higher financial system leverage mechanically magnifies the impact of given losses on bank net worth and credit supply. But secondly, and more importantly, elevated financial system leverage under shadow banking stimulates financial institutions to hold more risky assets. Banks in our economy face a portfolio choice between safe loans to non-financial firms and investments in a risky ‘bubble asset’ which offers a superior return to loans when the bubble survives but zero recovery when the bubble collapses. Intuitively, the attraction of the risky bubble to banks is greater when the profits from traditional loans (measured by the lending-deposit spread) are low.

Increases in the size of the shadow banking sector and the corporate bond market both create an environment of excess loan supply, resulting in lower spreads and a higher share of bank balance sheets that are allocated to the bubbly asset. But shadow banking has a larger impact because it gives banks greater access to leverage and boosts their profitability for any given level of loan spreads. Higher bank profits expand credit availability further, and lead to even greater excess supply of bank loans and to even more downward pressure on spreads.

In contrast, the corporate bond market has a small or even negative impact on bank leverage during the boom. This helps to achieve equilibrium in the loan market with a smaller decline in lending-deposit spreads. Consequently, the expansion of the corporate bond market leads to a less pronounced increase in bank risk-taking.

---

Shadow banking allows the financial sector to offer more loans with their existing bank capital. The corporate bond market allows borrowers to bypass banks to some extent, resulting in reduced bank loan demand.

Since borrower net worth is predetermined, overall credit demand cannot increase in the short run, resulting in excess bank loan supply in both situations.

This is because banks face binding capital constraints in our economy. All profits are retained and used to fund further credit.
Finally, the real impact of financial crises is greater under ‘shadow banking’ because the economy remains fully reliant on bank loan supply. Credit from the shadow banking sector depends on the financial health of sponsoring financial institutions, and therefore it collapses together with credit supply from the traditional financial sector. In contrast, the corporate bond market helps to insulate firms from fluctuations in bank credit supply as highlighted by Greenspan (1999). While banks contract loan supply in a crisis, the bond market is unaffected, helping to stabilize firms’ access to credit. So the growth of bank-independent sources of finance makes a banking crisis more likely but less costly.

The paper is structured as follows. Section 2 discusses the related literature, Section 3 provides some motivating observations based on US data. Sections 4 and 5 introduce the baseline economic environment with non-bank credit and describe the bubbly equilibrium of the economy. Section 6 goes in more detail into the conditions under which banks hold the bubble asset and become exposed to the bubble’s collapse. Section 7 uses numerical simulations to show how our two types of financial innovations increase banking sector fragility. Finally, section 8 concludes.

2 Related Literature

The literature on shadow banking has grown rapidly over the past several years. Gennaioli et al. (2012) focus on the role of the shadow banking system in creating quasi-safe assets. They stress the systemic fragility that can arise when risks from rare events are neglected. Our contribution is to analyze the impact of shadow banking on bank risk-taking and financial fragility in a framework with fully rational expectations. Alessandri et al. (2014) focus on the way losses by highly leveraged shadow banks can amplify macroeconomic shocks. Our treatment of shadow banking emphasizes high leverage similarly to Alessandri et al. (2014). However, in our
framework shadow banking arises due to incentives for regulatory arbitrage.\footnote{There has been substantial work on the shadow banking system in the finance literature. For example, Gorton (2010) argues that the shadow banking system undertakes maturity transformation without access to lender of last resort or deposit insurance. This left it vulnerable to classic bank runs. In addition, many authors (for example Pozsar et al. (2012) and Gai et al. (2011)) have stressed the complex and inter-connected nature of the shadow banking system and the dangers that this poses for systemic risk. While these are undoubtedly very important issues, we do not tackle them in this paper but leave their integration into our framework for future research.} This is motivated by the compelling empirical evidence in Acharya et al. (2013) that regulatory arbitrage stimulated the growth of Asset Backed Commercial Paper (ABCP) conduits during the 2000s. The balance sheets of these conduits were effectively insured by sponsoring banks but without the associated capital charge that a traditional financial institution would have had to make.

Interest in the corporate bond market has also grown since the Lehman crisis. Adrian et al. (2013), De Fiore and Uhlig (2015) and Gertler and Karadi (2013) show that the presence of corporate bond finance acted like a ‘spare tire’ for the corporate sector during the financial crisis. De Fiore and Uhlig (2011) study the determinants of the split between bank and corporate bond finance and argue that risky firms go to banks to take advantage of their superior monitoring skills while safer ones go to the bond market due to the cheaper funding it offers. Our framework also stresses banks’ superior monitoring skill similarly to De Fiore and Uhlig (2011) but does so in a simpler way. We assume that borrowers can pledge more collateral to banks compared to ordinary buyers of corporate bonds. This generates similar implications - bond finance is cheaper than bank finance and provides some funding diversification in the case of negative shocks to the banking sector. Uniquely in the literature, our paper examines the impact of the corporate bond market on banks’ incentives to hold risky assets.

Our paper is also related to the growing literature on rational bubbles in models of credit frictions (Caballero and Krishnamurthy (2006), Kocherlakota (2009), Martin and Ventura (2012), Farhi and Tirole (2012) and Aoki and Nikolov (2015) among others) because we model the ‘risky asset’ as arising from an asset price bubble which
can burst stochastically and deliver large losses to its holder. In our previous work (Aoki and Nikolov (2015)) we show that, when the banking system holds assets whose price contains a bubble component, this creates the possibility of large banking losses and a credit crunch when the bubble eventually collapses. In Aoki and Nikolov (2015) we relied primarily on moral hazard driven by government guarantees for banks in order to generate bank exposure to the asset price bubble. Without a financial safety net, banks buy few risky assets in equilibrium because their traditional safe activities are very profitable in an environment of binding bank balance sheet constraints. In this work, we show how substantial bank exposures to the bubble can arise from financial innovations that reduce banks’ profitability and ‘franchise values’.

3 Financial Innovations and Bank Profitability in the US

In this section we document a number of facts about the US financial system over the past thirty years. The key message we want to convey is that traditional commercial banks have found themselves increasingly competing with other finance providers since 1980. Figure 1 shows how the corporate bond market has grown relative to commercial bank credit since the Second World War. Despite recent volatility in the size of the outstanding stock of corporate bonds relative to bank loans, we can see that it has risen by 10-15 percentage points since 1980.

The corporate bond market expanded credit supply to high grade corporates and this market became increasingly difficult for banks to compete in, leading them to move into hitherto under-developed real estate lending. Figure 2 shows how, in the 1980s, banks switched their portfolios away from commercial and industrial loans (which fell from 40% to 17% of total loans) and towards real estate loans (which rose from 25% to around 60% of total loans).
Banks, however, continued to face competition from further financial innovation. Securitization started to grow in earnest around 1990 as shown by the expanding balance sheets of ABS issuers and broker-dealers (Figure 3). The growth of these ‘shadow bank’ entities was especially rapid after 2003. For example ABS issuers’ balance sheets expanded from around 30% of commercial bank assets in 2003 to 45% of bank assets in 2007, before collapsing back to 25% during the crisis. The growth of ABS increased the competition banks faced in the mortgage market which by then had become their largest source of lending business.

Such rapid financial innovation, de-regulation and growth in competition affected banks’ profitability. FDIC data presented in Figure 4 sheds more light on banks’ profitability over this period. Net interest income declined as a percentage of bank equity from a peak of over 50% in the late 1980s to 25% in recent years. At the same time, banks started to lend to more risky borrowers as evidenced by the growing loss provisions. After accounting for loss provisions, net interest income peaked in 1980 at 50% and declined sharply to almost 10% during the crisis. One important factor which allowed banks to maintain profitability is non-interest income which grew until 2000 although it has been declining as a share of bank equity since then.

This brief look at the US financial sector since 1980 reveals two key phenomena about the interaction of the banking system and non-bank sources of funds. First of all, the commercial banking sector has experienced an increasingly competitive environment and this has forced it to adapt by shifting towards real estate financing and relying increasingly on non-interest income. Secondly, even though these shifts in business models allowed banks to maintain profitability, this came at the expense of higher risk as evidenced by the volatile equity returns (Figure 5) and higher loss provisions. Furthermore, the data highlight the fact that the supply of these different sources of external finance have not all proved durable under stressed financial conditions. While the corporate bond market increased the funds it provided to the real economy, the ABS sector contracted strongly since 2007.
In the rest of this paper, we outline a model environment which can link these phenomena in a consistent framework. We show that the decline in bank profitability and the increase in bank risk-taking can be naturally linked to the growth in non-bank funding sources observed in the data. Finally, we argue that the durability (or lack thereof) of non-bank credit sources in times of crisis can be usefully linked to the dependence of these funding sources on the banking system itself.

4 The Model

The main aim of this paper is to study the way various financial innovations affect the risks to financial stability posed by fragile exuberance in financial and credit markets. We model fragile exuberance as the stochastic collapse of a rational bubble of the Samuelson-Tirole type and examine how the real economy reacts.

The bubble asset we will refer to frequently from now on is a durable but intrinsically useless asset which is in fixed aggregate supply. Its fundamental value is zero but, as we show in Aoki and Nikolov (2015), under credit frictions, there exist bubbly equilibria in which expectations can be coordinated on a positive valuation. The rational bubble is therefore a convenient and tractable means of introducing a risky asset whose price can fluctuate violently. As the bank exposure to the bubble will play a major role in generating real effects from the bubble’s collapse, much of our modelling efforts will concentrate on understanding the conditions under which banks undertake significant exposure to this risky asset.

The economy is populated with two main kinds of agents: entrepreneurs with heterogeneous productivity and banks. Low productivity entrepreneurs become savers and high productivity entrepreneurs become borrowers in equilibrium. Borrowing and lending can be intermediated by banks but also by two kinds of direct funding sources. One is the corporate bond market in which savers lend to borrowers without any role for banks. The other is the shadow banking system in which savers

---

5 We also discuss the condition in Section 6 of this paper.
lend directly to borrowers but pay fees to banks in exchange for guarantees that loan repayment will be enforced in the event of borrower default.

4.1 Entrepreneurs

We start by characterising the entrepreneurial sector. Each entrepreneur is endowed with a constant returns to scale production function which converts labor $h_t$ into output in the next period $y_{t+1}$.

$$y_{t+1} = a_i^t h_t,$$  \hspace{1cm} (1)

where $a_i^t$ is a productivity parameter which is known at time $t$.

In each period some firms are productive ($a_i^t = a^H$) and the others are unproductive ($a_i^t = a^L < a^H$). A productive entrepreneur in this period may become unproductive in the next period with probability $\delta$, and an unproductive entrepreneur in this period may become productive with probability $n\delta$. This probability is independent across entrepreneurs and over time. This Markov process implies that the fraction of productive entrepreneurs is stationary over time and equal to $n/(1 + n)$, given that the economy starts with such population distribution.\(^6\)

Entrepreneurs are ex-ante identical and have log utility over consumption streams

$$U^E = E_0 \sum_{t=0}^{\infty} \beta^t \ln c_t$$  \hspace{1cm} (2)

They purchase consumption $c_t$, bubbles $m_t^c$ at price $\mu_t$, they borrow amount $b_t^l$ from banks (negative $b_t^l$ means that the household places deposits with banks) and they borrow $b_t^m$ directly from savers. They also pay wages $w_t h_t$ to the workers they hire in order to receive future revenues $a_i^t h_t$ which the government taxes at rate $\tau_t$ after deducting debt repayments. $w_t$ and $h_t$ denote real wage and labor respectively.

\(^6\)We assume that the probability of the productivity shifts is not too large ($\delta + n\delta < 1$). This assumption implies that the productivity of each agent is persistent.
The flow of funds constraint is given by

\[ c_t + w_t h_t + m_t^x \mu_t - b_t^m - b_t^l = \]
\[ (1 - \tau_t) \left( a^t h_{t-1} - R_t^b b_{t-1}^l - \tilde{R}_t^b b_{t-1}^m + m_{t-1}^s \mu_t \right) \equiv (1 - \tau_t) z_t \quad (3) \]

where \( z_t \) stands for entrepreneur's net worth. \( R_t^b \) is equal to the loan rate \( R_t^l \) when the entrepreneur is a borrower and to the deposit rate \( R_t^d \) when he/she is a saver. \( \tilde{R}_t^b \) is the cost (return) of direct market loans. The cost for borrowers is \( \tilde{R}_t^b \) and the return to savers is \( \tilde{R}_t^d \). These returns may differ from \( R_t^m \) (the interest rate on the market bonds) if savers need to pay fees to banks in order to guarantee the repayment of the bonds. The relationship between these different interest rates is given by:

\[ \tilde{R}_t^d = R_t^m - (1 - \psi) p_t \quad (4) \]

where \( \psi \) is an index function which takes the value of 1 when savers can enforce marketable debt themselves and 0 when banks have a monopoly on debt enforcement and savers need to purchase enforcement guarantees from them in order to ensure they are repaid. \( p_t \) is the cost of bank guarantees per unit of the bond.\(^7\) We will discuss these market arrangements in some detail below. We will also use extensively the \( \psi \) notation as a way of examining the different reasons for directly traded debt with the minimum of repetitive derivations.

4.2 Loan Market Arrangements

In this paper we are interested in the interaction of non-bank funding sources with the banking system and consider the impact of two innovations that allow direct

\(^7\) Without a loss of generality, we assume that the savers do not pay the premium upfront but when the repayment with interest is received. This assumption is not crucial for our results. It is however very convenient because it ensures the timing of interest income and non-interest income is the same. This is what will ensure that the ‘shadow bank’ economy has an equivalent real allocation to a pure bank intermediation economy with a looser balance sheet constraint - a result which will be established later on in the paper.
lending between savers and borrowers.

4.2.1 The Corporate Bond Market

In one case, savers can directly enforce loans without the involvement of financial institutions. We will refer to this case as the ‘corporate bond market’. This means that $\psi = 1$ in equation (4) and no fees are paid to banks.

Banks remain in existence despite the presence of the corporate bond market because they have superior loan enforcement skills compared to savers. More specifically, we assume that entrepreneurs with expected revenues from production and bubble holdings\(^8\) $E_t (y_{t+1} + m^e_t \mu_{t+1})$ can pledge up to $\theta E_t (y_{t+1} + m^e_t \mu_{t+1})$ to banks but only up to $\theta (1 - \chi) E_t (y_{t+1} + m^e_t \mu_{t+1})$ to savers.\(^9\) Hence firms face two collateral constraints in the corporate bond economy. One limits what they can borrow from the market

$$E_t \Upsilon_{t+1} + R^m_t b^m_t \leq \theta (1 - \chi) E_t (y_{t+1} + m^e_t \mu_{t+1}),$$

where $R^m_t b^m_t$ is the promised repayment to corporate bond investors, and $E_t \Upsilon_{t+1}$ stands for expected tax payments.\(^10\)

The other constraint limits their total borrowing to the amount that can be pledged to savers.

---

\(^8\)We include bubble holdings in the collateral constraint only for completeness. In equilibrium, borrowing agents will not find it optimal to hold any other asset apart from productive projects. The proof of this statement is in Appendix C. The intuition is that, as long as bubbles are not completely collateralisable and require a downpayment ($\theta < 1$), they take up valuable net worth which could be used to expand production. In equilibrium, the return from production is higher for productive agents than the expected return on bubbles. Hence, high productivity entrepreneurs concentrate on leveraged production and do not hold the bubble asset.

\(^9\)Of course these pledgable fractions are not additive. For example, if $\theta (1 - \chi)$ fraction of firm revenues is directly pledged to savers, only $\theta \chi$ can be pledged to banks. The maximum fraction of the firm that can be pledged to all borrowers is $\theta$.

\(^10\)We assume that the tax authorities have a first call on collateral while private creditors are second in line. Hence under limited commitment, tax payments

$$E_t \Upsilon_{t+1} \equiv E_t \tau_{t+1} (\alpha^i h_t - R^i_t b_t + m^e_t \mu_{t+1})$$

crowd out the ability to borrow from private creditors.
enforced by banks with their superior enforcement technology

\[ E_t Y_{t+1} + R_t^m b_t^m + R_t^l b_t^l \leq \theta E_t \left( y_{t+1} + m_t \mu_{t+1} \right). \]  

(6)

This assumption is a simple way of capturing the intuition that some loans can be intermediated via a market - these are relatively safe loans whose enforcement requires little special skill. Other loans, however require the enforcement skills of bankers. Our framework will ensure that banks and markets coexist in equilibrium even though it will turn out that bank loans will be a more expensive means of debt finance than corporate bonds.\(^{11}\)

### 4.2.2 Shadow Banking

In another alternative case, which we will refer to as ‘shadow banking’ throughout the paper, we assume that even though some loans can be directly traded between borrowers and savers without necessarily going on banks’ balance sheets, the enforcement of these loans requires special skills only bankers have. This means that \( \psi = 0 \) in equation (4) - savers cannot enforce loan repayment by themselves. The only relevant constraint for entrepreneurs is, therefore, (6) which states that both market and bank loans are limited by banks’ ability to collect debts from entrepreneurs.

Due to the lack of any saver enforcement power, direct loans will only trade with the help of a bank guarantee which costs \( p_t \) per unit of bonds and, without a loss of generality, we assume that the savers pay the premium in the following period.\(^{12}\)

\(^{11}\)Our set up in which bank and bond finance co-exist within the same firm is consistent with the variable investment scale version of the Holmstrom and Tirole (1997) model.

In reality, bank and bond finance do not always co-exist within the same firm. Young and small firms borrow mostly from banks while old and large firms borrow more from capital markets. In contrast, our firms use bond and bank finance in the same proportions regardless of size or age.

In a separate exercise which is available upon request, we develop a model in which firms gain access to the bond market only as they get older. When they are young, they are bank-dependent. We show that, as long as old and young firms do not differ in their average productivities, the choice of whether to model bank and bond finance as co-existing within the same firm does not matter for aggregate dynamics.

\(^{12}\)This assumption is not crucial for our results. It is however very convenient because it ensures the timing of interest income and non-interest income is the same. This is what ensures that the ‘shadow bank’ economy has an equivalent allocation to a pure bank intermediation economy with
Our assumption on bank guarantees is inspired by the description in Acharya et al. (2013) of the way bank liquidity lines to ABCP vehicles were structured prior to the crisis. In our model, this guarantee is a promise to purchase the bonds from savers and enforce repayment in the event that the borrower threatens to default. In subsequent analysis we will demonstrate that the only reason such a market will exist is in order to provide avenues for regulatory arbitrage.

4.3 Banks

Bankers are risk neutral and live for a stochastic length of time. Once bankers receive an “end of life” shock, they liquidate all their asset holdings and consume their net worth before exiting. This shock hits with probability $1 - \gamma$. Banks maximize the following objective function:

$$U^B = E_0 \sum_{t=0}^{\infty} (\beta \gamma)^t c_t^b$$

subject to the following constraints explained below.

In each period the bank has net worth ($n_t$). It collects deposits ($d_t$) from the savers. Then it lends to the borrowers ($b_t$), purchases bubbles in non-negative quantities ($\mu_t m_t^b \geq 0$), issues guarantees ($s_t$) on market debt or consumes ($c_t^b$). We assume that intermediation is costless. The bank’s balance sheet is given by

$$c_t^b + b_t + \mu_t m_t^b = n_t + d_t.$$  

Crucially the bank issues guarantees $s_t$ off its balance sheet. This is why $s_t$ is not in the balance sheet constraint (8) although the fee income from guarantees ($p_t s_t$) appears in the net worth evolution equation. When the bubble does not burst, net

\footnote{In order to keep the population of bankers constant, we also assume that in each period measure $1 - \gamma$ of new bankers are born with small initial endowments (net worth). Since those small initial endowments do not affect the subsequent analysis, we do not analyze those explicitly.}
worth evolves according to:

\[ n_{t+1} = R_t^l b_t + \mu_{t+1} m_t^b - R_t^d d_t + p_t s_t \]  \hspace{1cm} (9)

while if it does burst it is given by:

\[ n_{t+1} = R_t^l b_t + \rho m_t^b - R_t^d d_t + p_t s_t. \] \hspace{1cm} (10)

The parameter \( \rho \) in equation (10) represents the government’s bailout policy. We assume that the government guarantees fraction \( \rho \) of the banks’ bubble investment in the event of a bubble collapse. This is a simple means of capturing the explicit or implicit guarantees given by the government to the banking system. In Aoki and Nikolov (2015) we show that we can micro-found this parameter using a model with banks under limited liability and deposit insurance. The bailout fraction \( \rho \) is a reduced-form representation of the deposit insurance subsidy which arises when banks can default leaving the government to pay the cost of compensating insured depositors. Since we have enriched the model substantially by adding non-bank funding sources, here we take the reduced form approach to bailouts in order to keep the analysis manageable.

Following Gertler and Karadi (2011), we model banks subject to limited commitment. More specifically, the banker may divert \( 1 - \lambda^m \) fraction of total liabilities. Once he diverts, a banker consumes the funds and closes his bank, remaining inactive until his ‘death’. Savers will therefore restrict the amount they deposit with the intermediary, according to the following incentive compatibility (borrowing) constraint which states that the value of diverted funds is less than or equal to the continuation value of the bank \( V (n_t) \) which will be defined later on in section 5:

\[ (1 - \lambda^m) (s_t + d_t) \leq V (n_t). \]  \hspace{1cm} (11)
The liabilities on the left-hand side include both on-balance-sheet deposits \((d_t)\) as well as off-balance-sheet guarantees \((s_t)\) which turn into divertable deposits as soon as the guarantee is called in.\(^{14}\) So for the purposes of the bank’s market-imposed borrowing constraint, on and off-balance-sheet liabilities require the same amount of bank capital.

In addition to the market borrowing constraint \((11)\), the bank also faces a regulatory constraint\(^{15}\) which is specified in terms of traditional bank liabilities \(d_t\) only:\(^{16}\)

\[
(1 - \lambda^*)d_t \leq V(n_t). \tag{12}
\]

The bank maximizes \((7)\) subject to \((8), (9), (10), (11)\) and \((12)\).

### 4.4 Workers

Unlike the entrepreneurs, the workers do not have access to the production technology nor any collateralizable asset in order to borrow. They maximize

\[
U^W = E_0 \sum_{t=0}^{\infty} \beta^t \left( c^w_t - \frac{h_t^{1+\eta}}{1 + \eta} \right) \tag{13}
\]

subject to their flow-of-funds constraint

\(^{14}\)We are not trying to model rigorously the microfoundations for the borrowing constraint \((11)\) but we have the following environment in mind.

Within every period, savers move first and divide their savings into conventional bank deposits and direct market loans to borrowers (backed by bank guarantees). Immediately after the bond market closes, borrowers have the opportunity to default on the bonds they just sold to savers. At this point, the deposit market can reopen in order to provide banks with funds with which to purchase the defaulted bonds from savers and enforce repayment.

After the closure of the deposit market, banks have the opportunity to divert deposits, which at this point could potentially include bank guarantees as well as conventional deposits. This explains why the left hand side of \((11)\) is specified in terms of \(s_t + d_t\).

\(^{15}\)Regulatory constraints are usually specified in terms of accounting definitions of equity:

\[
(1 - \lambda^*)d_t \leq n_t
\]

In a note available upon request we show that all our results continue to hold if we consider a regulatory capital constraint in terms of accounting definitions of equity rather than the market value of equity.

\(^{16}\)Here we assume that the regulatory constraint holds only ‘ex ante’. If the guarantees are called, the bank has to increase its balance sheet size beyond what is allowed by regulators. For simplicity (but also realistically) we assume that regulators exercise forbearance in such a situation and allow the bank to violate its capital adequacy ratios at least for one period.
\[ c_t^w + m_t^w \mu_t - b_t^w = w_t h_t + m_{t-1}^w \mu_t - R_{t-1}^d b_{t-1}^w. \]  

(14)

here superscript 'w' stands for 'workers'.

4.5  The Government

We assume that the only role for the government in this economy is to levy taxes on entrepreneurs and bail out the banking system when it makes losses. We assume that the government follows a balanced budget rule and does not issue government debt. Consequently taxes are only levied whenever bailout spending is necessary. In our framework this only happens when the bubble bursts. So when \( \mu_t \) suddenly collapses to zero, taxes satisfy the following condition

\[ \tau_t Z_t = \rho m_{t-1}^b \]

where \( Z_t \) is the aggregate wealth of entrepreneurs (defined in Appendix A) and \( m_{t-1}^b \) are the banks’ bubble holdings whose market value has collapsed to zero. For the rest of the time \( \tau_t = 0 \).

5  Equilibrium

Following Weil (1987) we consider a stochastic bubble that persists with probability \( \pi \). With probability \( 1 - \pi \) the bubble bursts and its value collapses to zero. We assume this probability is constant over time. Also, we assume that once bubbles burst they never arise again.

In equilibrium, due to the difference in their productivity, productive entrepreneurs borrow and unproductive entrepreneurs save. We focus on equilibria in which the productive entrepreneurs borrow up to their borrowing constraint.\(^{17}\)

\(^{17}\)This happens when the borrowing constraints are tight enough. See Aoki et al. (2009).
5.1 Entrepreneurs’ optimal behavior

Since the period utility function is logarithmic and there is no labour income or transfer income, entrepreneurs consume a constant fraction of net worth ($z_t$)

$$c_t = (1 - \beta) z_t,$$

(15)

and save the remaining $\beta$ fraction.

5.1.1 High productivity entrepreneurs

High productivity entrepreneurs enjoy better returns on production so they are the ones who borrow in equilibrium. We focus on an equilibrium in which the overall borrowing constraint (6) binds. When $R^l_t > R^m_t$ (which will be true in the case of the corporate bond market), entrepreneurs prefer to borrow as much as they can from corporate bond investors before going to banks for loans. Therefore constraint (5) binds too.

In the shadow banking economy savers cannot enforce debts and bank and market loans will be perfect substitutes in equilibrium.\footnote{This result will be shown in Section 5.2} Hence $R^l_t = R^m_t$ and borrowers will be indifferent as to whether they borrow from banks or from the market.

Since high productivity entrepreneurs earn a high rate of return on productive projects, they will find it unattractive to hold any other assets including the bubble asset. In other words, the value of holding a unit of the productive asset exceeds the value of holding a unit of the bubble asset:

$$E_t \left[ 1 - \frac{\tau_{t+1} \bar{R}^H_{t+1}}{c^H_{t+1}} \right] < E_t \left[ 1 - \frac{\tau_{t+1}}{c_{t+1}} \right] \frac{a^H}{w_t},$$

(16)

where $1/c^H_{t+1}$ represents the shadow value of wealth at time $t+1$ of the entrepreneur who is productive at time $t$\footnote{Namely, it is given by $c^H_{t+1} = (1 - \beta)Z^H_{t+1}$}, where expectation operator is taken over whether
bubble survives or crashes. $\tilde{\mu}_{t+1}$ represents the entrepreneur’s state-contingent payoff of bubble holdings:

$$\tilde{\mu}_{t+1} = \begin{cases} 
\mu_{t+1} & \text{with probability } \pi \\
0 & \text{with probability } 1 - \pi 
\end{cases}$$

(17)

where $\mu_{t+1}$ is the market value of the bubble on survival. We confirm throughout any numerical simulation of the model that (16) always holds in equilibrium. Intuitively, we will show in the next subsection that low productivity entrepreneurs will hold the bubble asset and therefore its expected return will follow closely their expected return on production ($a^L/w_t$). When $a^H > a^L$, high productivity agents will find it unattractive to hold the bubble under such conditions. When high productivity entrepreneurs invest only in productive projects, their rate of return on wealth ($r(a^H)$) is given by:\footnote{This equation is a compact way of representing $r(a^H)$ in a way that holds in the corporate bond as well as in the shadow bank economy. (18) is written as it is in the corporate bond economy whereas, strictly speaking, in the shadow bank economy, the equation is}

$$r(a^H) = \frac{a^H(1 - \tilde{\theta}_t)}{w_t - a^H \tilde{\theta}_t \left( (1 - \chi) / R^m_t + \chi / R^l_t \right)} \geq R^l_t.$$  

(18)

where $\tilde{\theta}_t$ is the maximum amount that can be pledged to private creditors after taking into account expected tax payments. The entrepreneur finances $(1 - \chi)$ fraction of her debts from the bond market and $\chi$ fraction from banks. Thus the denominator of (18) is the required downpayment for the unit labor cost which can be financed with market bonds and bank loans.

Furthermore, as long as $\tilde{\theta}_t < 1$, using the bubble asset as collateral is not attractive because the borrowing constraint forces the leveraged holder to use some of its own funds in purchasing the bubble. We show this result more formally in where $Z^H_{t+1}$ is given by equation (A.8).
Appendix C. Intuitively speaking, holding the bubble asset as collateral is costly for productive agents because it reduces the amount they can invest in their very high yielding productive projects. In subsequent sections, we first guess that productive agents hold no bubbles and verify it in any numerical solutions of the model.

When the productive entrepreneur only invests in production, her investment (employment) is given by

\[ h_t = \frac{\beta z_t}{w_t - a^H \theta_t \left( \left(1 - \chi \right) / R_m^t + \chi / R_l^t \right)}. \]  

(19)

The entrepreneur saves a \( \beta \) fraction of wealth \( z_t \) and uses her entire savings as a downpayment for wage payments to the workers she hires.

5.1.2 Low productivity entrepreneurs

Next, we characterize the optimal behavior of the low productivity entrepreneurs, who are savers in equilibrium. They can make bank deposits or buy market bonds and since both are safe, their rates of return (net of fees) must be equalized in equilibrium:

\[ R_d^t = R_m^t - (1 - \psi) p_t. \]

We can see easily that when \( \psi = 1 \) (i.e., when savers can enforce debts by themselves), \( R_d^t = R_m^t \). When the bank’s balance sheet constraint binds and \( R_d^t > R_l^t \), corporate bond funding is cheaper than bank loans as previously conjectured.

Low productivity entrepreneurs have two other means of savings: unleveraged production and investing in bubbles. When

\[ R_d^t > \frac{a^L}{w_t} \]

(20)

low productivity agents are inactive in production. However when the credit constraints on banks and borrowing entrepreneurs are tight enough, the productive
entrepreneurs cannot absorb all national saving and the low productivity technology may be viable in equilibrium. In this case:

\[ R_t^d = \frac{a^L}{w_t}. \]  

(21)

Bubbles are risky. When savers invest in bubbles as well as deposits, the arbitrage condition for bubbles is determined by the savers’ state-contingent wealth valuation

\[ E_t \left[ \frac{1 - \tau_{t+1}}{c_{t+1}^L} \tilde{\mu}_{t+1} \right] = E_t \left[ \frac{1 - \tau_{t+1}}{c_{t+1}^L} \mu_t \right] R_t^d, \]  

(22)

where \( 1/c_{t+1}^L \) represents the shadow value of wealth at time \( t + 1 \) of the entrepreneur who is unproductive at time \( t \). \(^{21}\)

As we showed in Aoki and Nikolov (2015), savers are the natural bubble holders when \( R_t^d > R_t^i \) and \( a^H > a^L \). Intuitively, when \( a^H > a^L \), high productivity entrepreneurs have an advantage in investing in productive projects while, when \( R_t^d > R_t^i \), banks will have a strong incentive to invest in loans. In contrast, savers have the worst investment opportunities (since \( a^L < a^H \)), which are dynamically inefficient for a large set of parameter values.

### 5.2 Banks’ optimal behavior

Next, we characterize the optimal behavior of a representative bank in our economy. The problem of the bank can be represented in recursive form as follows:

\[ V(n_t) = \max_{c_t^b, d_t, b_t, m_{t+1}, s_t} \left\{ c_t^b + \beta E_t \left[ \gamma V(n_{t+1}) + (1 - \gamma) n_{t+1} \right] \right\} \]  

(23)

\( V(n_t) \) is the value of a bank with net worth \( n_t \) which chooses current consumption, deposits, bubbles and loans optimally. This value is equal to current consumption

\[^{21}\)Namely, it is given by \( 1/c_{t+1}^L = (1 - \beta)Z_{t+1}^L \) where \( Z_{t+1}^L \) is given by equation (A.9).\]
and the expected future discounted value of bank net worth \( \beta E_t [\gamma V (n_{t+1}) + (1 - \gamma) n_{t+1}] \).

This value includes the continuation value of being a banker - this happens only if the banker survives with probability \( \gamma \). With probability \( 1 - \gamma \), the banker receives the end-of-life shock and consumes his entire net worth in the following period.

When \( \psi = 1 \) in equation (4) and savers do not need banks to enforce market bond repayment, banks’ guarantees will be worthless (\( p_t = 0 \)) and no guarantees will be issued (\( s_t = 0 \)). Then the banks have only on-balance-sheet liabilities and their borrowing constraint is given by the tighter of the two constraints (11) and (12):

\[
d_t \leq \min \left( \frac{1}{1 - \lambda^m}, \frac{1}{1 - \lambda^r} \right) V (n_t)
\]

Here we assume that the regulatory constraint is tighter so in the economy where \( \psi = 1 \), banks operate with the leverage desired by regulators.

When \( \psi = 0 \), banks are needed in the enforcement of market bonds and this gives financial institutions the possibility of having off-balance sheet liabilities \( s_t \). Since both debt guarantees and normal deposits attract the same market capital requirement, equilibrium requires that the bank charges a guarantee ‘insurance premium’ which is equal to the lending-deposit rate spread:

\[
p_t = R^d_t - R^d_t.
\]

(24)

The bank is then indifferent between lending to firms directly on-balance-sheet or indirectly off-balance-sheet. (11) becomes the only real balance sheet constraint for banks. Bankers choose the size of the shadow banking system (here given by guarantees \( s_t \)) to make sure that the regulatory capital constraint (12) is satisfied. This is given by the following condition:

\[
s_t = \left( \frac{1}{1 - \lambda^m} - \frac{1}{1 - \lambda^r} \right) V (n_t).
\]

(25)
When \( s_t \) is set using (25), (12) no longer affects bank total leverage and aggregate lending. In other words, if unregulated shadow banks are set up in the model, regulation becomes a ‘pure veil’.\(^{22}\) It influences only the split between the regulated and unregulated sectors without affecting the real allocations in the economy.\(^{23}\) This stark result captures well the spirit of the discussion of Tucker (2010) as well as the empirical evidence presented in Acharya et al. (2013). Faced with a strong profit motive for circumvention, the financial system will find ways in which to get around capital regulation regimes that focus on a narrow regulatory perimeter.

To summarize, the effect of shadow banking on financial intermediaries is to modify their collateral constraint in the following way:

\[
\left(1 - \hat{\lambda}\right) \hat{d}_t \leq V(n_t) \tag{27}
\]

where

\[
\hat{\lambda} = \psi \lambda^r + (1 - \psi) \lambda^m \tag{28}
\]

and

\[
\hat{d}_t = d_t + s_t. \tag{29}
\]

Note that \( s = 0 \) when \( \psi = 1 \) as discussed above. Hence, the ability to issue off-balance sheet guarantees is equivalent to relaxing the bank’s leverage constraint by \(^{22}\)This result depends on the assumption that insurance premia are paid in arrears rather than in advance. Without such an assumption, the exact equivalence will change but the spirit of the analysis will be preserved. \(^{23}\)We assumed for simplicity that the regulatory constraint takes the same form as the constraint imposed by markets. It may be more natural to assume that the regulatory constraint puts a constant upper bound on the bank’s book leverage such as:

\[
(1 - \lambda')d_t \leq n_t \tag{26}
\]

Even in this case, it continues to be true that borrowers are indifferent between borrowing from the traditional banking sector and from the shadow banking sector because the enforcement power is identical between the two. Also, the savers are indifferent between making deposits and buying the bank guaranteed bonds. Then, equation (24) holds and thus banks are indifferent between lending to firms directly on-balance sheet or indirectly off-balance sheet. Therefore the subsequent analysis will continue to hold even if we assume a more realistic regulation given by (26).
increasing \( \hat{\lambda} \) from \( \lambda^r \) to \( \lambda^m \).\(^{24}\)

Because of risk neutrality, we can guess that the value of the bank is a linear function of net worth \( n_t \)

\[
V(n_t) = \phi_t n_t
\]

(30)

When \( R^d_t > R^l_t \), the balance sheet constraint (27) binds and consumption is postponed until death.

Our guess for \( V(n_t) \) together with the bank’s value function (23) imply that the bank chooses bubble and loan holdings according to the following first order condition:

\[
E_t \left[ (1 - \gamma + \gamma \phi_t) \frac{\hat{\mu}_b t + 1}{\mu_t} \right] \leq E_t \left[ (1 - \gamma + \gamma \phi_t) \right] R^d_t.
\]

(31)

Here \( \hat{\mu}_b t + 1 \) represents the bank’s state-contingent payoff of bubble holdings:

\[
\hat{\mu}_b t + 1 = \begin{cases} 
\mu_{t+1} & \text{with probability } \pi \\
\rho \mu_t & \text{with probability } 1 - \pi
\end{cases}
\]

(32)

where \( \rho \) is a parameter which governs the (explicit or implicit) financial safety net for banks provided by the government. When the bubble bursts, banks receive a fraction \( \rho \) of their original bubble investment. The expectation operator is again taken over the bubble surviving or not. If equation (31) holds with strict inequality, then the bank will not invest in bubbles \( (m^b_t = 0) \). When the bank invests in bubbles (31) must hold with equality.

Finally, by substituting (11), (28), (30) and (31) into (23), \( \phi_t \) satisfies

\[
\phi_t = \frac{\beta E_t \left[ (1 - \gamma) + \gamma E_t \phi_{t+1} \right] R^l_t}{1 - \beta E_t \left[ (1 - \gamma) + \gamma E_t \phi_{t+1} \right] \frac{R^d - R^l}{\hat{\lambda}}}
\]

(33)

This expression states that the value of a unit of net worth for a banker is equal

\(^{24}\)In contrast, the corporate bond market does not affect the value of \( \hat{\lambda} \).
to the value of the returns on its loan book (the numerator), suitably boosted by leverage (the denominator). It also shows that $\phi_t$ is an increasing function of current and future spreads $R^l_t - R^d_t$.

### 5.3 Workers’ optimal behavior

Workers are risk-neutral and only wish to save when

$$R^d_t \geq \beta^{-1}. \tag{34}$$

Ours is a limited commitment economy, and we can guess (and later verify) that $R^d_t < \beta^{-1}$ at all times. Because the workers cannot operate the production technology, they cannot pledge collateral to lenders and cannot borrow. Hence our workers are hand-to-mouth consumers whose labour supply $h^s_t$ is given by

$$h^s_t = w_t^{1/\eta}. \tag{35}$$

### 5.4 Aggregation and market clearing

In the interest of brevity, the full set of aggregate equilibrium conditions is given in Appendix A.

### 6 The Determinants of Bank Risk-Taking

The bubble is the only risky asset in our model economy.\textsuperscript{25} Therefore, bank risk-taking manifests itself in higher bubble holdings by banks. In Aoki and Nikolov (2015) we showed that asset price bubbles pose significant risks to economic and

\textsuperscript{25}In general, rational bubbles circulate when they are affordable and attractive. Affordability means that bubbles must not grow too fast. In the case of deterministic bubble, the gross rate of return of bubbles must not be higher than the gross rate of economic growth which is unity in our model. Aoki and Nikolov (2015) shows that this happens when $\theta$ is at an intermediate level. Conditions for attractiveness are discussed in detail in this Section.
financial stability only when they are held by the banking system. In this section, we discuss bank risk-taking via bubble purchases. This provides a link to our earlier work but also highlights the way in which the financial innovations introduced in this paper affect bank risk-taking and financial fragility. The current section will show that the lending-deposit spread \((R^d_t - R^l_t)\) plays a major role in determining whether the bank is exposed to the bubble or not. This discussion is leading up to our simulation results in the next section where we demonstrate that corporate bond financing and shadow banking affect banks’ risk taking differently due to their differential impact on the lending-deposit spread.

To understand banks’ incentive to hold the bubble, we start with a risk-premium representation of the arbitrage condition for savers’ bubble holdings (equation (22)):

\[
E_t \left[ \frac{\tilde{\mu} t + 1}{\mu t} \right] = R^l_t - \frac{cov_t \left[ \frac{1 - \tau_{t+1}}{c_{t+1}}, \tilde{\mu} t + 1 \right]}{E_t \left[ \frac{1 - \tau_{t+1}}{c_{t+1}} \right]} \equiv R^l_t + \kappa^l_t
\]

where \(\kappa^l_t\) is the risk premium required by savers in order to hold the risky bubble.

When banks hold the bubble, the following condition should hold in equilibrium (this time based on equation (31)):

\[
E_t \left[ \frac{\hat{\mu} t + 1}{\mu t} \right] = R^l_t - \frac{cov_t \left[ (1 - \gamma + \gamma \phi_{t+1}), \frac{\hat{\mu} t + 1}{\mu t} \right]}{E_t \left[ (1 - \gamma + \gamma \phi_{t+1}) \right]} \equiv R^l_t + \kappa^B_t
\]

where \(\kappa^B_t\) is the risk premium required by bankers.

It may seem surprising that the bankers who have linear period utility would require a risk premium in order to hold the bubble. This is due to the effect of binding borrowing constraints and the high and time varying value of internal funds that this brings about. The shadow value of wealth for bankers is given by \((1 - \gamma + \gamma \phi_{t+1})\) — a term whose time series behavior is dominated by \(\phi_{t+1}\) — the value of wealth.
for surviving bankers. Binding borrowing constraints for banks lead to an aggregate shortage of bank capital and a lending spread \((R^l > R^d)\). Then, being a banker is valuable, because it allows the enjoyment of the super-normal profits created by these spreads. \(\phi_{t+1}\) is the net present value of these super-normal profits - the ‘franchise value’ of the bank.

Bankers would like to transfer wealth to states of the world in which these super-normal profits are high - i.e. where lending spreads and \(\phi_{t+1}\) are relatively high. We show later that spreads in our model are high when bank capital is depleted by bubble collapse. This makes banks risk-averse with respect to gambles that affect aggregate banking system net worth and hence the profits of normal banking activities. In the case of the bubble asset, the risk aversion of banks depends on whether the financial system as a whole is exposed to the bubble’s bursting or not.\(^{26}\)

The expected return for banks also includes the anticipated bailout payment from the government even when the bubble bursts. Since we assume that the savers do not have access to bailout policy, we have

\[
E_t \left[ \frac{\tilde{\mu}_{t+1}^b}{\mu_t} \right] = E_t \left[ \frac{\tilde{\mu}_{t+1}^b}{\mu_t} \right] + (1 - \pi) \rho,
\]

where \((1 - \pi) \rho\) is the expected value of any bailouts that accrue to a bank holding the bubble. This means that equation (37) above can be alternatively expressed as:

\[
E_t \left[ \frac{\tilde{\mu}_{t+1}^b}{\mu_t} \right] + (1 - \pi) \rho = R^l_t + z^R_t.
\]  

(38)

Banks hold the bubble if the expected bubble return (including any bailout assistance) is higher than their alternative use of funds (the loan rate) and the risk premium required by the bank.

Substituting the savers’ condition for holding bubbles into the banks’ condition

\(^{26}\)This is the ‘Last Bank Standing’ effect highlighted by Perotti and Suarez (2002).
we get the following condition for banks holding the bubble:

\[(1 - \pi) \rho + \gamma_t^L - \gamma_t^B = R_t^l - R_t^d.\] (39)

The left-hand side of the above equation consists of the expected bailout term \((1 - \pi) \rho\) and the relative risk premium term \((\gamma_t^L - \gamma_t^B)\) whose determinants we discussed earlier in this section.

The lending spread \((R_t^l - R_t^d)\) term on the right-hand side of (39) is the ‘bank franchise value’ channel we focus on in this paper. The presence of binding borrowing constraints on banks introduces a wedge between lending and deposit rates and allows banks to earn rents on the loans they extend to the real economy. When the balance sheet constraint binds, this lending-deposit spread is high and buying the risky asset is less attractive to the bank. Its access to intermediation opportunities allows it to earn large profits with default-free loans only.

In the next section we use numerical simulations to show that financial innovation and shadow banking both increase competition in the financial sector and drive down banks’ lending spreads and franchise values. Holding all the terms on the left hand side of (39) fixed, this makes banks more willing to hold the bubble asset and increases the risks that a bubble collapse poses to banks’ financial health. We find, however, that shadow banking reduces spreads and increases bank bubble holdings by a larger amount due to the fact that it expands aggregate financial system leverage further.

7 Non-Bank Credit and Financial Fragility

In this section we answer the central question of our paper: how does growing competition from different non-bank finance sources affect bank and macro-economic fragility? To do this we compare the impact of the two financial innovations discussed in the paper - the growth of the corporate bond market and the growth of
the shadow banking sector. Following our approach in Aoki and Nikolov (2015) the possibility of a financial crisis arises due to the presence of an asset price bubble, which can burst and damage the net worth of those economic agents who hold it.²⁷

We discuss the properties of the model based on numerical simulations. The model calibration follows the methodology of Aoki and Nikolov (2015) and full details are available in Appendix B of this paper.

In order to study the effects of financial innovations on financial fragility, we take an economy which already has a bubble worth 10% of GDP and this bubble is endogenously held by savers with bank holdings at zero. This provides a natural benchmark of an economy which is subject to substantial financial market risk but whose banks are perfectly safe. We then hit this economy with two additional financial innovation shocks in order to see whether these additional shocks stimulate banks to buy the bubble asset.

The first shock expands the ability of savers to lend directly to borrowers without the involvement of banks. This is our ‘corporate bond market’ scenario and it is implemented in our model by a decrease in $\chi$ to 0.846 from its baseline value of unity. The second shock allows banks to shift a quarter of their assets off-balance sheet through the mechanism described in Section 5. Both scenarios are calibrated so as to generate an increase in non-bank funding sources equal to 25% of commercial bank loans. This is the approximate increase in ABS between 2000 and 2006 as shown in Figure 3.

Our main interest is in the following questions. How does the creation of these new lending institutions affect existing financial intermediaries’ incentives to take risk by holding bubbles? How does this affect the size of the downturn in the real economy when the bubble finally bursts?

²⁷Notice that emergence of direct finance between savers and borrowers does not increase the overall pledgeability of borrowers’ assets. In both the corporate bond and shadow bank economy, total pledgeability is still given by the overall collateral parameter $\theta$. Both mitigate frictions in financial intermediation. However, corporate bond finance does not rely on bank capital while the shadow banking sector (implicitly or explicitly) relies on bank capital. As is shown below this difference is crucial.
Figures 6 and 7 compare the evolution of the economy under the ‘shadow bank’ and ‘corporate bond’ scenarios. In those figures, the financial innovations occur in period 2 of the simulations while the bubble collapses in period 5. Since our model is calibrated to the annual frequency, this implies that the boom lasts 3 years. All the numbers in the chart are scaled by the baseline in which the economy is in a bubbly equilibrium but no financial innovation shocks have occurred. Figure 6 focuses on bank balance sheet variables, and Figure 7 displays the evolution of bank net worth, credit supply and real output during our simulations.

Figure 6 shows that the boom-time impact of the two innovations on bank leverage is very different. Shadow banking expands leverage by around a quarter relative to baseline while the corporate bond market decreases leverage during the boom period. The reason for this difference lies in the fact that the expansion of corporate bond market encourages the growth of unleveraged direct loans from savers to borrowers. This decreases demand for bank loans, squeezing bank profitability and reducing bank franchise values \( \phi_t \).\(^{28}\) Since the bank’s collateral constraint depends on its charter value (a market value concept), this reduces leverage. Note that bank profitability also declines under the shadow banking simulation and this, ceteris paribus, decreases bank leverage too. However, the direct impact of the ability to shift assets off balance sheet (where they are subjected to less stringent market based capital requirements) dominates and consolidated financial system leverage expands.

This difference in bank leverage is reflected in the evolution of lending spreads. Bank lending margins \( (R^l_t - R^d_t) \) decline in both simulations but under shadow banking the decline is greater. To see why this is, first note that the higher leverage from shadow banking has a direct expansionary effect on credit supply when the

\(^{28}\)Recall from (33) that \( \phi_t \) is an increasing function of current and future spreads \( R^l_t - R^d_t \).
bank’s borrowing constraint binds.\textsuperscript{29} This can be seen from the expression below

\[ L_s^t = \left( 1 + \frac{\phi_t}{1 - \lambda} \right) \gamma N_t \]  

(40)

where \( L_s^t \) is loan supply. Since loan demand is constrained by the (unchanged) net worth of the entrepreneurs, the lending deposit spread falls in order to reduce bank net worth over time and keep credit supply aligned with credit demand.

But leverage has an additional positive effect on credit supply though its impact on bank profitability when \( R_l^t - R_d^t > 0 \). Consider the expression for the growth rate of aggregate bank equity:

\[ \frac{N_{t+1}}{N_t} = \gamma \left[ R_l^t + \frac{\phi_t}{1 - \lambda} (R_l^t - R_d^t) \right] \]  

(41)

When the bank’s leverage \( \left( \frac{\phi_t}{1 - \lambda} \right) \) increases, this makes bank net worth grow faster in equilibrium, expanding credit supply. As a result, credit spreads have to fall even further in order to offset the impact of higher leverage on bank profitability and achieve loan market equilibrium.

In contrast, when the corporate bond market expands, bank loan supply is initially unaffected but demand for bank loans falls as borrowers switch to the bond market. Hence, similarly to the case of an expansion in the shadow banking sector, intermediary net worth has to shrink in order for the market for bank loans to clear. This is again achieved via reduced spreads. However, the growth of the corporate bond market causes a decline in bank leverage (driven by a lower \( \phi_t \)) and this helps the adjustment in bank net worth to take place with a smaller fall in spreads. As a result, banks increase risky asset holdings by less compared to the shadow bank economy.

As the last panel in Figure 6 shows, the different behavior of bank leverage and

\textsuperscript{29}The increase in leverage arises due to the higher value of \( \lambda \). \( \phi_t \) actually declines because lending spreads fall and this has the effect of lowering leverage. However, the direct effect of higher \( \lambda \) dominates and \( \phi_t / (1 - \lambda) \) increases.
spreads has a significant impact on banks’ holdings of bubbly assets. As discussed in Section 6, lower lending spreads makes holding bubbles more attractive for banks and therefore bank risk-taking increases more during the shadow bank simulation compared to the corporate bond one. This occurs both through greater holdings of risky (bubble) assets but also through higher leverage which makes bank net worth more sensitive to potential losses. Finally, in Figure 7 we can see that output and lending expand significantly during the boom before the bubble finally bursts in period 5 of the simulation.

The collapse of the bubble can be seen clearly in Figure 6. The value of the bubble relative to GDP goes from just under 10% of annual output to zero in period 5. In the ‘shadow banking’ case, financial sector losses lead to a significant decline in credit when the bubble crashes. In contrast, the expansion of the corporate bond market largely offsets the impact of the bursting of the bubble and, as a result, lending experiences a very small decline. Bank net worth collapses under both scenarios but the impact of the credit crunch is more pronounced under shadow banking both in terms of higher spreads and in terms of lower credit quantities.

Looking at the evolution of output under the two scenarios, it is clear that the post-crisis growth slowdown is more pronounced in the case of shadow banking. This is, first of all, because of the more severe banking crisis and larger increase in spreads but also because the higher lending spreads affect all corporate liabilities. In contrast, the corporate bond market provides funding \((1 - \chi)\) fraction of all loans\) whose cost is independent of the health of banks and does not increase during the crisis. This is the ‘spare tire’ effect of the corporate bond market and it limits the deleveraging required by the corporate sector following the collapse of the bubble.

In the above simulations we assumed that the shadow banking sector survived intact following the collapse of the bubble. In reality, the Lehman crisis led to a tightening of the capital rules that govern shadow banking entities, resulting in most off-balance-sheet exposures moving back on to banks’ balance sheets (where they
are subject to capital charges). Figures 8 and 9 show such a scenario by assuming that the shadow banking sector is closed following the bubble collapse. In contrast, the corporate bond market continues to operate.

The closure of the shadow banking system hits the economy with a permanent decline in bank leverage in addition to the net worth loss from the bubble collapse. As a result, output and credit undergo a deep and permanent fall in the simulations shown in Figures 8 and 9. Bank net worth actually recovers much faster on the back of an enormous increase in credit spreads which go to almost 4pp in the crisis period, helping to ‘recapitalize’ the banking system. However, the ‘recapitalization’ is done at the expense of entrepreneurs who suffer high interest rates and lower net worth, contracting investment and output in the process.

8 Conclusions

This paper asks the question of how the growth of non-bank financing sources affects systemic risk. We find that the answer crucially depends on whether the availability and cost of these non-bank sources of finance is truly independent of the health of the banking sector. We contrast two deliberately extreme cases: one in which the non-financial sector improves its ability to enforce debts and another in which the banking sector shifts bank loans off balance sheet in order to circumvent regulation. In both cases, non-bank financing grows but only in the first case this growth can be sustained regardless of the level of bank capital.

Both of the shocks to the supply of non-bank credit sources we consider expand aggregate credit supply. This increases investment and output but also lowers lending-deposit spreads for banks, reducing their franchise values. As a result, bubbly assets become more attractive for banks and they increase their exposure to them. The banking system is vulnerable to a collapse in the bubble and risks to bank capital are high in both cases.
Where the two experiments differ is in the consequences of bank losses for the real economy. In the ‘corporate bond’ scenario, the fall in bank capital triggers only a very mild credit crunch because the corporate sector has an alternative source of funds to go to. This is the ‘spare tire’ highlighted by Greenspan (1999) — diversification of funding sources makes banks less important in providing funds to the real economy and this diminishes the costs of systemic banking crises. In contrast, the ‘shadow banking’ scenario features a more severe credit crunch.
Appendices

A Aggregate Equilibrium Conditions

In this Appendix we set up the aggregate equilibrium conditions. Let the total supply of the bubble asset be normalized to 1. Then, the market clearing condition for bubble is given by

$$m^e_t + m^b_t = 1,$$

where $m^e_t$ and $m^b_t$, respectively, denote the shares of the bubble held by unproductive entrepreneurs and banks.

Let $Z^H_t$ and $Z^L_t$, respectively, denote aggregate wealth of the productive and unproductive entrepreneurs. Then we can characterize the aggregate equilibrium as follows. From (19) the aggregate employment of the productive entrepreneurs is given by:

$$H^H_t = \frac{\beta Z^H_t}{w_t - a^H \theta_t (1 - \chi)/R^m_t + \chi/R^l_t)}.$$  

When (21) holds, the unproductive entrepreneurs are indifferent between making deposits and producing, thus their aggregate saving is split as follows

$$H^L_t = \beta Z^L_t - D_t - B^M_t - m^L_t \mu_t$$

where $D_t$ and $B^M_t$, respectively, denote aggregate deposits and aggregate directly traded loans.

Let us turn to banks. In the shadow banking economy, total loan supply is determined by banks’ market borrowing constraint. Hence the aggregate quantities of deposits and bank guarantees are given by:

$$S_t + D_t = \frac{\phi_t}{(1 - \lambda^m)} \gamma N_t.$$  

Equation (A.2) present the aggregate condition from the corporate bond economy. It also holds (trivially) in the shadow bank economy because $R^m_t = R^l_t$.  

34
where $S_t$ denotes the aggregate off-balance sheet guarantees issued by banks. The split between on and off-balance-sheet activities is given by the regulatory constraint:\footnote{Notice that $1 - \gamma$ fraction of banks exits in each period and consumes their net worth. Therefore the aggregate net worth of the operating banks is given by $\gamma N_t$. See, also, footnote 13}

$$D_t \leqslant \frac{\phi_t}{(1 - \lambda^r)\gamma N_t}$$  \hspace{1cm} (A.5)

Finally the size of the direct loan market is given by the supply of bank guarantees ($S_t = B_t^M$) and the aggregate balance sheet of the operating banks is given by

$$D_t + \gamma N_t = B_t^L + m_t^L \mu_t.$$  \hspace{1cm} (A.6)

In the corporate bond economy, $S_t = 0$ and the size of the corporate bond market is determined by savers’ loan enforcement abilities:

$$R_t^d B_t^M = \theta (1 - \chi) a^H H_t^H$$  \hspace{1cm} (A.7)

The bank’s balance sheet constraint is determined by regulation: in other words (A.5) holds while (A.6) does not.

Unproductive entrepreneurs become productive in the next period with probability $n\delta$ and productive entrepreneurs continue to be productive with probability $1 - \delta$. Their rates of return are given by (18) and (20). Therefore the net worth of the productive and unproductive entrepreneurs evolve according to the following two conditions:

$$Z_{t+1}^H = (1-\delta) \frac{a^H (1 - \tilde{\theta}_t)}{w_t - a^H \tilde{\theta}_t \left( (1 - \chi) / R_t^M + \chi / R_t^d \right)} \beta Z_t^H + n\delta \left[ R_t^d \left( \beta Z_t^L - m_t^e \mu_t \right) + m_t^e \mu_{t+1} \right].$$  \hspace{1cm} (A.8)

$$Z_{t+1}^L = \delta \frac{a^H (1 - \tilde{\theta}_t)}{w_t - a^H \tilde{\theta}_t \left( (1 - \chi) / R_t^M + \chi / R_t^d \right)} \beta Z_t^H + (1-n\delta) \left[ R_t^d \left( \beta Z_t^L - m_t^e \mu_t \right) + m_t^e \mu_{t+1} \right].$$  \hspace{1cm} (A.9)
Aggregate output is given by

\[ Y_t = a^H H_{t-1}^H + a^L H_{t-1}^L. \] (A.10)

Finally, aggregate bank net worth is given by

\[ N_{t+1} = \gamma \left[ R_t^d B_t^l + m_t^b \mu_{t+1} - R_t^d D_t + (1 - \psi) \left( R_t^l - R_t^d \right) S_t \right]. \] (A.11)

The markets for goods, labour, capital, bubble, loan and deposit must clear. Goods market clearing implies that aggregate saving must equal to aggregate investment.

\[ \beta (Z_t^H + Z_t^L) + \gamma N_t = w(H_t^H + H_t^L) + \mu_t. \] (A.12)

From (35), labour market clearing implies

\[ w_t^\eta = H_t^H + H_t^L. \] (A.13)

Equations (21), (22), (31), (33), (A.1)-(A.13) jointly determine 17 variables \( R_t^d, R_t^l, w_t, H_t^H, H_t^L, Y_t, \phi_t, D_t, S_t, B_t^L, B_t^M, Z_t^H, Z_t^L, N_{t+1}, \mu_t, m_t^e, m_t^b \) with three states \( Z_t^H, Z_t^L, N_t \).\(^{32}\)

**Definition 1** Competitive bubbly equilibrium without government is a sequence of decision rules \( \{ H_t^H, H_t^L, Y_t, D_t, S_t, B_t^L, B_t^M, m_t^e, m_t^b \}_{t=0}^{\infty} \), aggregate state variables \( \{ Z_{t+1}^H, Z_{t+1}^L, N_{t+1} \}_{t=0}^{\infty} \) and a price sequence \( \{ R_t^d, R_t^l, w_t, \phi_t, \mu_t \}_{t=0}^{\infty} \) such that: (i) entrepreneurs, banks and workers optimally choose decision rules \( \{ H_t^H, H_t^L, Y_t, D_t, S_t, B_t^L, B_t^M, m_t^e, m_t^b \}_{t=0}^{\infty} \) taking the evolution of aggregate states, prices and idiosyncratic productivity opportunities as given; (ii) the price sequence \( \{ R_t^d, R_t^l, w_t, \phi_t, \mu_t \}_{t=0}^{\infty} \) clears the goods, labor, capital, loan, bubble and deposit markets and (iii) the equilibrium evolution of state variables \( \{ Z_{t+1}^H, Z_{t+1}^L, N_{t+1} \}_{t=0}^{\infty} \) is consistent with the individual choices of entrepreneurs, banks and workers and with the exogenous evolution of productive

\(^{32}\)By Warlas law one of these equations is redundant.
opportunities at the individual firm level.

B Calibration

We follow Aoki and Nikolov (2015) for model calibration. We have 11 parameters \( \{ \eta, a^H/a^L, \rho, \delta, n, \theta, \chi, \gamma, \beta, \lambda^m, \lambda^r \} \). For simplicity we set \( \chi = 1 \) in the baseline calibration thus having no market financing in the baseline economy. There is little consensus regarding \( \eta \), the Frisch elasticity of labour supply. Micro-data evidence suggests a value close to zero based on the labour supply behavior of primary earners. At the other extreme, Gertler and Kiyotaki (2010) set the Frisch elasticity to 10, justifying their choice by the presence of labour market frictions. We pick a value of \( \eta = 5 \), which is within the range set in calibrating macro models. Following Aoki et al. (2009) we set a value for \( a^H/a^L = 1 \).

The value of \( \rho \) (the fraction of bubble holdings which is bailed out by the state) is set to 0.5. For simplicity we parameterize bailout guarantees in this parsimonious way but our reasoning is based on the model of deposit insurance under supervisory constraints developed in Aoki and Nikolov (2015). If banks enjoyed deposit insurance but faced no supervision, then it would be optimal for them to place all their bubble holdings in separate banking firms who would deliver zero recovery value in the bust. If banker equity was around 10% of total bank assets, this could justify a \( \rho \) closer to 0.9. Assuming some supervisory constraints in the fraction of the bubble in total bank assets still allows us to consider banks which hold 50-60% of their assets in bubbly securities. This is not an unreasonable number for some Irish banks or Spanish Cajas in recent years could justify our choice of \( \rho = 0.5 \) or even higher.

We calibrate the remaining 6 parameters to match the steady state predictions of the model in the absence of bubbles to 7 moments in the US data\(^{33}\). These are (1) the real loan rate minus the growth rate of real GDP and minus intermediation costs; (2) the real deposit rate minus the growth rate of real GDP; (3) commercial bank

\(^{33}\)More information on the data sources used in calibration are given in Table A1 in the Appendix.
leverage; (4) average corporate leverage; (5) average leverage for highly leveraged corporates; (6) the rate of return on bank equity and (7) the ratio of M2 to GDP.

Calibration targets (1) and (2) deserve further discussion. For simplicity, we assume no growth and no intermediation costs. Growth in the US has averaged close to 3% per annum since the second world war. We are interested in the dynamic efficiency of the investments of US savers and banks which is why we subtract the real growth rate from the real return on deposits and loans.

In addition, when it comes to evaluating the dynamic efficiency of banks’ loan investments, we need to take intermediation costs into account. FDIC data on US commercial banks’ cash flow sources reveals that there are substantial intermediation costs (80% of those are labour costs) and we assume that all of these arise due to loan issuance rather than deposit taking. This assumption is reasonable given the labour intensive nature of arranging loans, monitoring them and then recovering them if they become non-performing. We subtract these loan costs from banks’ real loan returns to get the final numbers shown in Table A2 below. Full details of data sources and construction are available in Table A1:
<table>
<thead>
<tr>
<th>Theor. concept</th>
<th>Data concept</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real bank loan rate</td>
<td>Real prime loan rate-GDP growth-costs</td>
<td>FRB, Table H.15, FDIC, BEA</td>
</tr>
<tr>
<td>Real deposit rate</td>
<td>Real M2 own rate - GDP growth</td>
<td>FRED</td>
</tr>
<tr>
<td>Expected inflation</td>
<td>Average CPI inflation (All Urban Consumers)</td>
<td>FRED</td>
</tr>
<tr>
<td>Expected real GDP growth</td>
<td>Average real GDP growth (chained measure)</td>
<td>FRED</td>
</tr>
<tr>
<td>Deposit stock</td>
<td>M2</td>
<td>FRED</td>
</tr>
<tr>
<td>Nominal GDP</td>
<td>Nominal GDP</td>
<td>FRED</td>
</tr>
<tr>
<td>Bank leverage</td>
<td>Bank Debt Liabilities/Bank Net Worth</td>
<td>FRB, Table H.8</td>
</tr>
<tr>
<td>Average corporate leverage</td>
<td>Corporate Debt/Corporate Net Worth</td>
<td>Welch (2004)</td>
</tr>
<tr>
<td>Leverage of indebted corporates</td>
<td>Debt/Corp Net Worth for high leverage corporates</td>
<td>Welch (2004)</td>
</tr>
<tr>
<td>Bank rate of return on equity</td>
<td>Bank rate of return on equity</td>
<td>FDIC</td>
</tr>
</tbody>
</table>
The parameter values that come out of our calibration exercise are given in Table A3 below.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>Probability of exiting a productive spell</td>
<td>0.177</td>
</tr>
<tr>
<td>$n$</td>
<td>Ratio of productive to unproductive agents</td>
<td>0.039</td>
</tr>
<tr>
<td>$a^H/a^L$</td>
<td>Relative productivity</td>
<td>1.100</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Bailout fraction for bubble holdings</td>
<td>0.500</td>
</tr>
<tr>
<td>$\eta$</td>
<td>Frisch elasticity of labour supply</td>
<td>5.000</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Entrepreneur’s collateral fraction</td>
<td>0.626</td>
</tr>
<tr>
<td>$\chi$</td>
<td>Disadvantage of savers in loan enforcement</td>
<td>1.000</td>
</tr>
<tr>
<td>$\lambda^r$</td>
<td>Regulatory capital constraint parameter</td>
<td>0.765</td>
</tr>
<tr>
<td>$\lambda^m$</td>
<td>Market capital constraint parameter</td>
<td>0.000</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Survival rate of bankers</td>
<td>0.867</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Entrepreneur’s discount rate</td>
<td>0.946</td>
</tr>
</tbody>
</table>
C Bubbles as Collateral

In the main body of the paper we stated that entrepreneurs do not want to hold the bubble asset in order to use it as collateral. Here we show that this is indeed the case. We compute the utility value of buying a bubble at price $\mu_t$, borrowing against it the maximum amount possible ($\tilde{\theta}_t E_t \mu_{t+1}/R^i_t$) and then using this borrowed amount as downpayment for a leveraged productive investment. Entrepreneurs prefer this strategy to simply undertaking a leveraged productive investment if the following condition holds:

$$E_t \left[ (1-\tau_{t+1}) \left( \frac{a^H (1-\tilde{\theta}_t)}{w_t - \theta_t a^H / R^i_t} \frac{\tilde{\theta}_t E_t \mu_{t+1}/\mu_t}{\mu_t} + \frac{\mu_{t+1} - \tilde{\theta}_t E_t \mu_{t+1}}{\mu_t} \right) \left( \frac{1}{c^H_{t+1}} \right) \right] \geq \frac{a^H (1-\tilde{\theta}_t)}{w_t - \theta_t a^H / R^i_t} E_t \left( \frac{1}{c^H_{t+1}} (1-\tau_{t+1}) \right).$$

(C.1)

When $a^H/a^L$ is high, (C.1) will not hold and productive entrepreneurs will not find it profitable to hold bubbles in order to use them as collateral. The intuition is that since $\tilde{\theta}_t < 1$, the collateral constraint forces the leveraged holder to use some of his own funds in purchasing the bubble. This is costly for productive agents because it reduces the amount they can invest in their very high yielding productive projects. When $a^H/a^L$ is high and $\tilde{\theta}_t$ is sufficiently below unity, productive agents do not hold bubbles.

In the steady state with safe bubbles ($\pi = 1$), this can be verified analytically. In this deterministic steady state, $\tilde{\theta}_t = \theta$, $\tau_t = 0$, and $\mu_{t+1}/\mu_t = 1$ at all times. Therefore equation (C.1) reduces to

$$\frac{a^H (1-\theta)}{w - \theta a^H / R^i} \left( \frac{\theta}{R^i} \right) \geq \frac{a^H (1-\theta)}{w - \theta a^H / R^i}.$$

(C.2)

Since $R^i > R^d = 1$, it is easily shown that (C.2) is violated.

---

$^34_{t+1} Z^H_t = (1-\beta) Z^H_{t+1}$ is consumption of the productive entrepreneur where $Z^H_{t+1}$ is given by equation (A.8).
References


Gertler, M., Karadi, P., 2013. Qe1 vs. 2 vs. 3: A framework to analyze large scale asset purchases as a monetary policy tool. International Journal of Central Banking 9 (S1), 5–53.


Figure 1: US corporate bond stock as % of US commercial bank loans (Source: US Flow of Funds)

Figure 2: Commercial & Industrial and Real Estate loans as a % of total US commercial bank loans (Source: FDIC)
Figure 3: Assets of ABS issuers and Broker-Dealers as a proportion of US commercial bank loans (Source: US Flow of Funds)

Figure 4: Sources of US Commercial bank profits (Source: FDIC)
Figure 5: US Commercial banks' net profit as a % of total equity (Source: FDIC)

Figure 6: Bank leverage, lending spreads and bubble holdings in the baseline
(All numbers apart from Bubble/GDP are relative to no-financial-innovation saver bubble baseline)
Figure 7: Bank net worth, output and credit supply in the baseline
(All numbers are relative to no-financial-innovation saver bubble baseline)

Figure 8: Bank leverage, lending spreads and bubble holdings
(shadow banking closed when the bubble bursts)
(All numbers apart from Bubble/GDP are relative to no-financial-innovation saver bubble baseline)
Figure 9: Bank net worth, output and credit supply in the baseline
(shadow banking closed when the bubble bursts)
(All numbers are relative to no-financial-innovation saver bubble baseline)