Search for Yield

David Martinez-Miera
Universidad Carlos III de Madrid

Rafael Repullo
CEMFI

July 2016

Abstract

We present a model of the connection between real interest rates, credit spreads, and the structure and the risk of the banking system. Banks intermediate between entrepreneurs and investors, and choose the monitoring intensity on entrepreneurs' projects. We first characterize the equilibrium for a fixed aggregate supply of savings, showing that safer entrepreneurs will be funded by nonmonitoring (shadow) banks and riskier entrepreneurs by monitoring (traditional) banks. We also show that a savings glut reduces interest rates and spreads, increases the relative size of the shadow banking system and the probability of failure of the traditional banks. The dynamic version of the model provides a framework for understanding the emergence of endogenous boom and bust cycles, as well as the procyclical nature of the shadow banking system, the existence of countercyclical risk premia, and the low levels of interest rates and spreads leading to the buildup of risks during booms.

JEL Classification: G21, G23, E44

Keywords: Savings glut, real interest rates, credit spreads, bank monitoring, shadow banks, financial stability, banking crises, boom and bust cycles.

This paper is based on Repullo's Walras-Bowley Lecture at the 2014 North American Summer Meeting of the Econometric Society. We are very grateful to Guillermo Caruana, Vicente Cuñat, Giovanni Dell'Ariccia, Pablo D'Erasmo, Charles Goodhart, Michael Gordy, Hendrik Hakenes, Juan Francisco Jimeno, Nobuhiro Kiyotaki, Michael Manove, Claudio Michelacci, Jean-Charles Rochet, Hyun Song Shin, and Javier Suarez for their valuable comments, and Dominic Cucic and Álvaro Remesal for their research assistance. Financial support from the Spanish Ministry of Economy and Competitiveness, Grants No. ECO2014-59262-P (Repullo) and ECO2013-42849-P (Martinez-Miera), and from Banco de España (Martinez-Miera) is gratefully acknowledged. E-mail: david.martinez@uc3m.es, repullo@cemfi.es.

1 Introduction

The connection between interest rates and financial stability has been the subject of extensive discussions and a significant amount of (mostly empirical) research. This paper contributes to this literature by constructing a theoretical model of the relationship between real interest rates, credit spreads, and the structure and the risk of the banking system. It thus provides a framework to understand how a "global savings glut" that reduces the level of long-term real interest rates, noted by Bernanke (2005) and Caballero, Fahri, and Gourinchas (2008), can generate incentives to "search for yield" and increases of risk-taking that can lead to financial instability, as noted by Rajan (2005) and Summers (2014).

The model shows that an increase in savings reduces interest rates and interest rate spreads, increases the relative size of the originate-to-distribute (shadow) banking system, and increases the probability of failure of the originate-to-hold (traditional) banks.[1] Moreover, the model generates endogenous boom and bust cycles: the accumulation of savings leads to a reduction in rates and spreads and an increase in risk-taking that eventually materializes in a bust, which reduces savings, starting again the process of wealth accumulation that leads to a boom. The model also yields a number of empirically relevant results such as the procyclical nature of risk-taking in the traditional banking system, the procyclicality of the shadow banking system, and the existence of countercyclical risk premia. These findings contribute to our understanding of the role of financial factors in economic fluctuations.

The paper starts with a simple partial equilibrium model of bank lending with three types of risk-neutral agents: entrepreneurs, investors, and a bank. Entrepreneurs seek bank finance for their risky investment projects. The bank, in turn, needs to raise funds from a set of (uninsured) investors. Following Holmström and Tirole (1997), the bank can monitor the entrepreneurs' projects at a cost, which reduces their probability of default. As monitoring is not contractible there is a moral hazard problem. We characterize the optimal contract

[1] Our use of the term shadow banking follows the Financial Stability Board (2014): "The shadow banking system can broadly be described as credit intermediation involving entities and activities outside of the regular banking system." They note that some authorities and market participants prefer to use other terms such as "market-based financing" instead of "shadow banking."

between the bank and the investors, showing that there are circumstances in which the bank chooses not to monitor entrepreneurs and others in which it chooses to monitor them. We associate the first case to (shadow) banks that originate-to-distribute, and the second case to (traditional) banks that originate-to-hold.

The partial equilibrium results show that which case obtains depends on the spread between the bank's lending rate and the expected return required by the investors, which under risk-neutrality equals the safe rate. In particular, a reduction in this spread reduces monitoring, and makes it more likely that the bank will find it optimal to originate-to-distribute.

To endogenize interest rates and interest rate spreads we embed our model of bank finance into a static general equilibrium model in which a large set of heterogeneous entrepreneurs, that differ in their observable risk type, seek funding for their investment projects from a competitive banking sector. Assuming that the higher the total investment in projects of a particular risk type the lower the return, we characterize the equilibrium for a fixed aggregate supply of savings, and show that safer entrepreneurs will borrow from originate-to-distribute banks while riskier entrepreneurs will borrow from originate-to-hold banks.

We then analyze the effects of an exogenous increase in the aggregate supply of savings, showing that it will lead to a reduction in interest rates and interest rate spreads, an increase in investment and in the size of banks' lending to all types of entrepreneurs, an expansion of the relative size of the shadow banking system, and a reduction in the monitoring intensity and hence an increase in the probability of failure of the traditional banks. These results provide a consistent explanation of a number of stylized facts of the period preceding the 2007-2009 financial crisis; see, for example, Brunnermeier (2009).

Although we focus on the effects of an exogenous increase in the supply of savings, the same effects obtain when there is an exogenous decrease in the demand for investment, due for example to a negative productivity shock. Thus, the model provides an explanation of the way in which changes leading to a reduction in the equilibrium real rate of interest, as those noted by Summers (2014), can be linked to an increase in financial instability.

Next we consider three extensions of our static model. First, we show that the effect of a savings glut on financial stability critically depends on the increase in the size of the traditional banks. When banks that originate-to-hold cannot increase their balance sheet (and adjust their loan rates), there will be a greater increase in the size of the shadow banking system, a greater reduction in the safe rate, and wider spreads for the traditional banks, so they will become safer. The assumption of a fixed size may be rationalized in terms of some capacity constraint that cannot be immediately relaxed. But as soon as originate-to-hold banks are able to expand they will become riskier. This result allows us to distinguish between the short- and the long-run effects of a savings glut, and provides a rationale for the idea that the buildup of risks in the traditional banking sector happens when (real) interest rates are "too-low for too-long."

The second extension deals with the case where investors are risk-averse. We show that a reduction in risk aversion has similar effects as a savings glut except for the level of the safe rate, which goes up instead of down. This provides a simple way to empirically distinguish a savings glut from a reduction in investors' risk appetite. The intuition is that when investors are less risk-averse, there is a shift in investment toward riskier entrepreneurs that reduces the funds available for safer ones. This leads to a reduction in loan rates for the former and an increase in loan rates for the latter, which reduces spreads and hence banks' monitoring incentives.

The third extension analyzes a variation of the model in which the riskiest entrepreneurs may not be able to fund their projects. In this setup, a savings glut will expand the set of (riskier) entrepreneurs that get funded.

Finally, we consider a dynamic version of the static model in which investors are infinitely lived and the aggregate supply of savings is endogenous. Specifically, the supply of savings at any date is the outcome of agents' decisions at the previous date together with the realization of a systematic risk factor that affects the return of entrepreneurs' projects. For good realizations of the risk factor, aggregate savings will accumulate leading to lower interest rates and spreads, which translate into higher risk-taking and a fragile financial system. In this situation the economy is especially vulnerable to a bad realization of the risk factor, which can lead to a crisis. The associated reduction in aggregate savings leads to higher interest rates and spreads, which translate into lower risk-taking and a safer financial system. Then savings will grow, restarting the process that leads to another boom. In this manner, the model generates endogenous boom and bust cycles.

The dynamic model yields other interesting and potentially testable results. First, interest rates and interest rate spreads are countercyclical. Second, during booms the safe rate may be below investors' subjective discount rate, and it may even be negative. Third, the shadow banking system is highly procyclical. Fourth, even though investors are risk-neutral, they behave as if they were risk-averse, so risky assets have positive risk premia. Fifth, even though investors' preferences do not change over time, such risk premia are countercyclical.

The brief review of the literature that follows discusses the relation to previous studies and the evidence on some of these predictions.

Literature review This paper is linked to different strands of the (theoretical and empirical) literature on the relationship between interest rates, financial frictions, the structure of the financial system, and the business cycle.

Our interest in the effects of financial frictions on macroeconomic activity relates to numerous studies following the seminal papers of Bernanke and Gertler (1989), Bernanke, Gertler, and Gilchrist (1996), and Kiyotaki and Moore (1997).[2] We have chosen to introduce these frictions using the moral hazard setup of Holmström and Tirole (1997). We depart from their model by focussing exclusively on the banks' moral hazard problem, endogeneizing the return structure that entrepreneurial projects offer in a competitive setup, and introducing heterogeneity in the ex-ante risk profile of entrepreneurs instead of in their net worth. In their characterization of equilibrium, entrepreneurs with low net worth borrow from monitoring banks while those with high net worth are directly funded by the market. In contrast, in our setup riskier entrepreneurs borrow from monitoring banks while safer entrepreneurs borrow from nonmonitoring banks, which could be interpreted as market finance.

[2] See Quadrini (2011) and Brunnermeier, Eisenbach, and Sannikov (2012) for surveys of macroeconomic models with financial frictions, and Adrian, Colla, and Shin (2013) for a review of the performance of these models in explaining key features of the 2007-2009 financial crisis.

Most papers that analyze the role of financial intermediaries in economic fluctuations focus on leverage; see, for example, Gertler and Kiyotaki (2010), Repullo and Suarez (2013), and Adrian and Shin (2014). We depart from this literature by considering a model in which banks have no equity capital. Our focus on the effect of endogenously determined interest rates on banks' decisions in a general equilibrium setting links our findings to those of Boissay, Collard, and Smets (2015). They analyze a model with an interbank market where lower interest rates make riskier banks more prone to borrow from safer banks. Their paper, like ours, generates endogenous boom and bust cycles which are driven by banks' strategic responses to changes in interest rates. But we ignore the interbank market, and focus on the effect of interest rates on banks' monitoring decisions.[3]

Our work is related to a large volume of research spurred following the 2007-2009 financial crisis. On the one hand, our paper provides a theoretical framework that links a savings glut with the level of interest rates and the increases in risk-taking noted by Rajan (2005) and Summers (2014) among many others. On the other hand, it yields some predictions regarding the behavior of interest rates and spreads, risk premia, and the structure and the risk of the banking system that are in line with recent empirical findings. For example, Lopez-Salido, Stein and Zakrajšek (2015) show that the widening of credit spreads following a period of low spreads is closely tied to a contraction in economic activity.[4] Our results on risk premia are also in line with Gilchrist and Zakrajšek (2012), who find a negative relationship between risk premia and economic activity, and Muir (2014), who finds that risk premia increase substantially in financial crises. Finally, our results on the procyclicality of shadow banking are consistent with the evidence in Pozsar, Adrian, Ashcraft, and Boesky (2012).

Many empirical papers analyzing the link between interest rates and banks' risk-taking focus on monetary policy. Although we have a real model without nominal frictions, some of this evidence is also in line with our predictions; see, for example, Jimenez, Ongena, Peydro, and Saurina (2014), Altunbas, Gambacorta, and Marques-Ibanez (2014), Dell'Ariccia,

[3]It should be noted that, as in Brunnermeier and Sannikov (2012) or He and Krishnamurthy (2012), we do not analyze a linearized version of the model but instead solve the full equilibrium dynamics.

[4]They interpret this result in behavioral terms (a change in "credit market sentiment"), whereas our story does not rely on changes in investors' preferences.

Laeven, and Suarez (2014), and Ioannidou, Ongena and Peydro (2015). Interestingly, our paper provides a rationale for the idea that low interest rates are dangerous from a financial stability perspective when they are low for a long period of time. However, our story is driven by the behavior of real interest rates, and is therefore not necessarily related to the stance of monetary policy.

Structure of the paper Section 2 presents the partial equilibrium model of bank finance under moral hazard. Section 3 embeds the partial equilibrium model into a general equilibrium setup, characterizing the equilibrium for a fixed aggregate supply of savings and analyzing the effects of an increase in the supply of savings. Section 4 considers three extensions of the general equilibrium model, which allow us to discuss the possible differences between the short- and long-run effects of a savings glut, the effect of having risk-averse instead of risk-neutral investors, and the way in which a savings glut can expand the set of (riskier) entrepreneurs that get funded. Section 5 analyzes the dynamic version of the model that generates endogenous booms and busts, and Section 6 concludes. The proofs of the analytical results are in the Appendix.

2 Partial Equilibrium

Consider an economy with two dates ($t = 0, 1$), a large set of potential *entrepreneurs*, a large set of risk-neutral *investors*, and a single risk-neutral *bank*. Entrepreneurs have investment projects that require external finance, which can only come from the bank. The bank, in turn, needs to raise funds from the investors, which are characterized by an infinitely elastic supply of funds at an expected return equal to R_0.

Each entrepreneur has a project that requires a unit investment at $t = 0$ and yields a stochastic return $\widetilde{R}$ at $t = 1$ given by

$$\widetilde{R} = \begin{cases} R, & \text{with probability } 1 - p + m, \\ 0, & \text{with probability } p - m, \end{cases} \tag{1}$$

where R and p are constant parameters, and $m \in [0, p]$ is a variable that captures the bank's monitoring intensity. Monitoring increases the probability of getting the high return R, but entails a cost $c(m)$. The monitoring cost function $c(m)$ satisfies $c(0) = c'(0) = 0$, $c'(m) \geq 0$, $c''(m) > 0$, and $c'''(m) \geq 0$. A special case that satisfies these assumptions and will be used for our numerical results is the quadratic function

$$c(m) = \frac{\gamma}{2} m^2, \tag{2}$$

where $\gamma > 0$. We assume that monitoring is not observed by the investors, so there is a moral hazard problem.

The bank can only fund a limited set of projects, taken to be just one for simplicity. Thus, entrepreneurs will be in the short side of the market and so they will only be able to borrow at the rate R that leaves them no surplus.

There are two possible modes of finance. The bank can keep the loan until maturity (*originate-to-hold*) or sell it to the investors (*originate-to-distribute*). We assume that the bank sells the loan when it is indifferent between keeping and selling it. Since monitoring is costly, and it is not observed by the investors, the bank will never monitor the entrepreneur when it is going to sell the loan, because it will get no compensation for its monitoring. Hence, originate-to-hold obtains when it is optimal for the bank to monitor the entrepreneur (i.e. set $m > 0$), and originate-to-distribute obtains when the bank prefers to do no monitoring (i.e. set $m = 0$).

To characterize the optimal mode of finance, suppose that the bank borrows from the investors at a rate B, chooses a monitoring intensity $m \in [0, p]$, and lends to the entrepreneur at the rate R.

An *optimal contract* between the bank and the investors is a pair (B^*, m^*) that solves

$$\max_{(B,m)} \left[(1 - p + m)(R - B) - c(m)\right] \tag{3}$$

subject to the bank's incentive compatibility constraint

$$m^* = \arg\max_{m} \left[(1 - p + m)(R - B^*) - c(m)\right], \tag{4}$$

the bank's participation constraint

$$(1 - p + m^*)(R - B^*) - c(m^*) \geq 0, \tag{5}$$

and the investors' participation constraint

$$(1 - p + m^*)B^* = R_0. \tag{6}$$

The incentive compatibility constraint (4) characterizes the bank's choice of monitoring m^* given the promised repayment B^*, and the participation constraints (5) and (6) ensure that the bank makes nonnegative profits, net of the monitoring cost, and that the investors get the required expected return on their investment.

An interior solution to (4) is characterized by the first-order condition

$$R - B^* - c'(m^*) = 0. \tag{7}$$

Solving for B^* in the participation constraint (6), substituting it into the first-order condition (7), and rearranging gives the equation

$$c'(m) + \frac{R_0}{1 - p + m} = R. \tag{8}$$

Since we have assumed $c'''(m) \geq 0$, the function in left-hand side of this equation is convex in m. Let $\underline{R}$ denote the minimum value of this function in the feasible range $[0, p]$, that is

$$\underline{R} = \min_{m \in [0,p]} \left(c'(m) + \frac{R_0}{1 - p + m} \right). \tag{9}$$

The following result shows the condition under which bank finance is feasible and characterizes the corresponding optimal contract between the bank and the investors.

Proposition 1 *Bank finance is feasible if $R \geq \underline{R}$, in which case the optimal contract between the bank and the investors is given by*

$$m^* = \max \left\{ m \in [0, p] \mid c'(m) + \frac{R_0}{1 - p + m} \leq R \right\} \text{ and } B^* = \frac{R_0}{1 - p + m^*}. \tag{10}$$

Proposition 1 shows that bank finance is feasible if the lending rate R is gretare than or equal to the minimum value $\underline{R}$ defined by (9), in which case the optimal contract is characterized by the highest value of m that satisfies

$$c'(m) + \frac{R_0}{1 - p + m} \leq R.$$

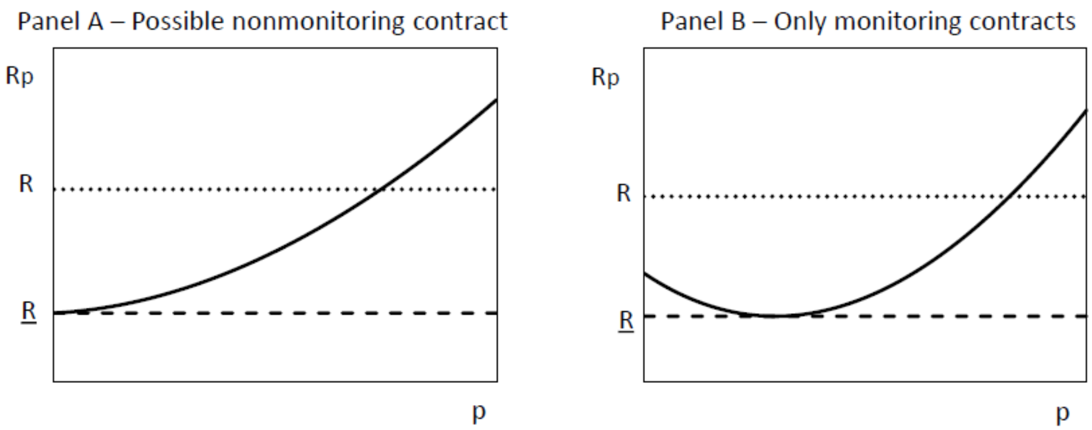

Figure 1. Characterization of the optimal contract

Panel A shows a case in which the optimal contract may entail zero monitoring (dashed line), and Panel B a case where the optimal contract always has positive monitoring.

Monitoring in the optimal contract may be at the corner with zero monitoring $m^* = 0$, at the corner with full monitoring $m^* = p$, or it may be interior $m^* \in (0, p)$. The first case corresponds to the originate-to-distribute mode of finance, while the second and third cases correspond to the originate-to-hold mode of finance.

Figure 1 illustrates the two modes of finance for the quadratic monitoring cost function. Panel A shows a case where the slope of the function in left-hand side of (8) is positive at the origin, in which case the optimal contract may entail $m^* = 0$ (for $R = \underline{R}$). Panel B shows a case where the slope of this function is negative at the origin, in which case the optimal contract always entails $m^* > 0$.

We next derive some interesting comparative static results on the optimal contract, assuming that it involves an interior level of monitoring.

Proposition 2 *If* $R > \underline{R}$ *we have*

$$\frac{\partial m^*}{\partial R_0} < 0 \text{ and } \frac{\partial m^*}{\partial R} > 0.$$

Thus, a reduction in the spread $R - R_0$, due to either an increase in the funding cost R_0 or a decrease in the lending rate R, reduces optimal monitoring, thereby increasing the bank's portfolio risk. For sufficiently low spreads, the bank may find it optimal to choose zero monitoring, switching from originate-to-hold to originate-to-distribute. Figure 1 illustrates the second result in Proposition 2: whenever bank finance is feasible, a reduction in the lending rate R (from the dotted to the dashed lines) always reduces monitoring m^*.

Summing up, we have presented a partial equilibrium model of bank finance under moral hazard that shows that *banks' monitoring incentives and hence banks' portfolio risk depends on the spread between lending and borrowing rates.* A reduction in the spread reduces monitoring, and makes it more likely that the bank will find it optimal to originate-to-distribute. However, the model assumes that interest rates are exogenous. To construct a model of the search for yield phenomenon we need to endogenize these rates, to which we turn now.

3 General Equilibrium

This section embeds our partial equilibrium model of bank finance into a general equilibrium setup in which a set of heterogeneous entrepreneurs seek finance for their risky projects. We characterize the equilibrium for a fixed aggregate supply of savings, showing that safer entrepreneurs will borrow from originate-to-distribute (shadow) banks while riskier entrepreneurs will borrow from originate-to-hold (traditional) banks. We then analyze the effects of an increase in the supply of savings, showing that it will lead to a reduction in interest rates and interest rate spreads, and an increase in the risk of the banking system. Finally, we consider whether the equilibrium is (constrained) inefficient, showing that a social planner subject to the same moral hazard problem as the banks would shift investments toward safer entrepreneurs.

Consider an economy with two dates ($t = 0, 1$) and a large set of potential entrepreneurs with observable types $p \in [0, 1]$. Entrepreneurs have investment projects that require external finance, which can only come from banks. Banks are risk-neutral agents that specialize in lending to specific types of entrepreneurs. To simplify the presentation, we will assume that for each type p there is a single bank that only lends to entrepreneurs of this type.[5] Banks, in

[5]Without loss of generality, we could have many banks lending to each type of entrepreneur. What might be restrictive is the assumption of bank specialization. However, Repullo and Suarez (2004) show that, for

turn, need to raise funds from a set of investors, which are characterized by a *fixed aggregate supply of savings w.*

Each entrepreneur of type p has a project that requires a unit investment at $t = 0$ and yields a stochastic return $\widetilde{R}_p$ at $t = 1$ given by

$$\widetilde{R}_p = \begin{cases} R_p, & \text{with probability } 1 - p + m, \\ 0, & \text{with probability } p - m, \end{cases} \tag{11}$$

where $m \in [0, p]$ is monitoring intensity of its bank. As before, monitoring is costly and the monitoring cost $c(m)$ satisfies our previous assumptions. The returns of the projects of entrepreneurs of each type p are assumed to be perfectly correlated (but we could have correlation across different types; see Section 5). This implies that the bank's return per unit of loans is identical to the individual project return, which is given by (11).

We assume that the success return R_p is a decreasing function $R(x_p)$ of the aggregate investment of entrepreneurs of type p, denoted x_p. Thus, the higher the aggregate investment x_p the lower the return R_p.

This assumption may be rationalized by assuming that entrepreneurs of type p produce (in case of success) output x_p which is an intermediate input sold at a price R_p to a set of final good producers. Specifically, suppose that for each p there is a continuum of final good producers with heterogeneous productivity θ_p, which is distributed according to the density $g(\theta_p) = a\sigma\,(\theta_p)^{-(\sigma+1)}$, where $a > 0$ and $\sigma > 1$.[6] Each producer can transform a unit of the intermediate input into θ_p units of the final good, which is assumed to have a unit price. For any price of the intermediate input R_p, only producers with productivity $\theta_p \geq R_p$ will operate. Hence, given a supply of the intermediate input x_p we must have

$$\int_{R_p}^{\infty} g(\theta_p)\; d\theta_p = a\,(R_p)^{-\sigma} = x_p, \tag{12}$$

which implies

$$R_p = R(x_p) = \left(\frac{x_p}{a}\right)^{-1/\sigma}. \tag{13}$$

---

a model with insured deposits, limited liability implies that it is optimal for banks to specialize.

[6] The assumption of firms with heterogeneous productivities follows the approach of Melitz (2003).

Notice that σ is the elasticity of the demand for the intermediate input, while a is a (proportional) demand shifter. This function (with $a = 1$) will be used to derive the numerical results of the paper.

Alternatively, we could introduce a representative consumer with a utility function over the continuum of goods produced by entrepreneurs of types $p \in [0,1]$. Specifically, we could assume that

$$U(q,x) = q + \frac{a\sigma}{\sigma - 1}\int_0^1 \left(\frac{x_p}{a}\right)^{\frac{\sigma-1}{\sigma}} dp, \tag{14}$$

where q is the consumption of a composite good, $x = \{x_p\}_{p\in[0,1]}$, $a > 0$ and $\sigma > 1$. Maximizing the utility of the representative consumer subject to the budget constraint

$$q + \int_0^1 R_p x_p \, dp = I$$

gives a first-order condition that also implies (13).

If the bank lending to entrepreneurs of type p sets a loan rate L_p, then a measure x_p of these entrepreneurs will enter the market until $L_p = R_p = R(x_p)$. Thus, as in the partial equilibrium setup, entrepreneurs will only be able to borrow at a rate that leaves them no surplus.

To determine equilibrium loan rates, we assume that the loan market is contestable. Thus, although there is a single bank that lends to each type, the incumbent could be undercut by an entrant if it were profitable for the entrant to do so.

The strategy for the analysis is going to be as follows. First, we characterize the investment allocation corresponding to any given safe rate R_0, which is derived from the condition that investors must be indifferent between funding banks lending to entrepreneurs of different types. Then we introduce the market clearing condition that equates the aggregate demand for investment to the aggregate supply of savings to determine the equilibrium safe rate R_0^*.

By contestability, a bank lending to entrepreneurs of type $p = 0$ will set a rate equal to the return R_0 of their projects, since at a lower rate it will make negative profits and at a higher rate it will be undercut by another bank. Similarly, banks lending to entrepreneurs of types $p > 0$ will set the lowest feasible rate, which by Proposition 1 (together with the perfect correlation assumption) is given by

$$\underline{R}_p = \min_{m\in[0,p]} \left(c'(m) + \frac{R_0}{1-p+m} \right). \tag{15}$$

The assumptions on the monitoring cost function $c(m)$ imply that we have a corner solution with zero monitoring if and only if

$$c''(0) - \frac{R_0}{(1-p)^2} \geq 0,$$

which gives $p \leq \widehat{p}$, where

$$\widehat{p} = 1 - \sqrt{\tfrac{R_0}{c''(0)}}. \tag{16}$$

Thus, banks lending to (safer) entrepreneurs of types $p \leq \widehat{p}$ will originate-to-distribute, and banks lending to (riskier) entrepreneurs of types $p > \widehat{p}$ will originate-to-hold. In what follows we will assume that $R_0 < c''(0)$, so $\widehat{p} \in (0,1)$.[7]

The intuition for this result is that since monitoring is especially useful for riskier entrepreneurs, they will have an incentive to borrow from originate-to-hold (monitoring) banks, and since monitoring is less useful for safer entrepreneurs (and useless for those with $p = 0$), they will borrow from originate-to-distribute (nonmonitoring) banks.

For banks that originate-to-distribute ($p \leq \widehat{p}$) loan rates are given by

$$R_p = \underline{R}_p = \frac{R_0}{1-p}, \tag{17}$$

where we have used the assumption $c'(0) = 0$. This result implies $(1-p)R_p = R_0$, so the expected return of the banks' investments equals the funding cost. Thus, profits in the originate-to-distribute mode of finance will always be zero.

For banks that originate-to-hold ($p > \widehat{p}$) loan rates are given by

$$R_p = \underline{R}_p = c'(m_p) + \frac{R_0}{1-p+m_p}, \tag{18}$$

where the monitoring intensity m_p satisfies the first-order condition[8]

$$c''(m_p) - \frac{R_0}{(1-p+m_p)^2} = 0. \tag{19}$$

[7] The model also works with $R_0 \geq c''(0)$, but in this case monitoring is so profitable that no bank (except the one lending to safe entrepreneurs of type $p = 0$) would originate-to-hold.

[8] Notice that we cannot have a corner solution with $m_p = p$ since the slope of the function in the right-hand-side of (15), evaluated at $m_p = p$, satisfies $c''(p) - R_0 \geq c''(0) - R_0 > 0$, where we have used $c'''(m) \geq 0$ and $R_0 < c''(0)$.

This result implies

$$(1-p+m_p)R_p - R_0 - c(m_p) = (1-p+m_p)c'(m_p) - c(m_p) > (1-p)c'(m_p) > 0,$$

where we have used (18) and the fact that $m_p c'(m_p) > c(m_p)$ by the convexity of the monitoring cost function. Thus, profits in the originate-to-distribute mode of finance will always be positive.

We can now state the following result.

Proposition 3 *For any given safe rate $R_0 < c''(0)$, there exists a marginal type $\widehat{p} \in (0,1)$ given by* (16) *such that banks lending to entrepreneurs of types $p \leq \widehat{p}$ will originate-to-distribute, and banks lending to entrepreneurs of types $p > \widehat{p}$ will originate-to-hold. Higher types p will be characterized by higher spreads $R_p - R_0$ and (for $p > \widehat{p}$) higher monitoring m_p. Moreover, increases in the safe rate R_0 will lead to a reduction in the marginal type $\widehat{p}$, an increase in spreads $R_p - R_0$, and (for $p > \widehat{p}$) an increase in monitoring m_p.*

Thus, for a given safe rate R_0, spreads and monitoring intensities are higher for riskier entrepreneurs, and they are both increasing in the safe rate.

We are now ready to define an equilibrium, which requires to specify the investment x_p of the different types of entrepreneurs, and hence the rates $R_p = R(x_p)$ at which they will borrow. By our previous results, both will be a function of the equilibrium safe rate R_0^*.

Formally, an *equilibrium* is an investment allocation $\left\{x_p^*\right\}_{p\in[0,1]}$ and corresponding loan interest rates $R_p^* = R(x_p^*)$ such that loan rates satisfy

$$R_p^* = \underline{R}_p^* = \min_{m\in[0,p]} \left(c'(m) + \frac{R_0^*}{1-p+m} \right), \text{ for all } p \in [0,1], \tag{20}$$

and the market clears

$$\int_0^1 x_p^* \, dp = w. \tag{21}$$

Condition (20) follows from the assumption that the market for lending to entrepreneurs of each type p is contestable, so equilibrium loan rates will be at the lowest feasible level $\underline{R}_p^*$ implied by the equilibrium safe rate R_0^*. Condition (21) ensures that the aggregate demand for investment is equal to the aggregate supply of savings w. Notice that the investors' participation constraint ensures that they all get the same expected return R_0^*, regardless of the type of bank they fund.

The equilibrium marginal type will be given by

$$p^* = 1 - \sqrt{\tfrac{R_0^*}{c''(0)}} \tag{22}$$

We will restrict attention to (sufficiently high) values of w so that $R_0^* < c''(0)$ and $p^* \in (0, 1)$.

3.1 An increase in the supply of savings

To analyze the effects of an exogenous increase in the supply of savings w notice that the market clearing condition (21) may be written as

$$F(R_0^*) = \int_0^1 R^{-1}(R_p^*)\ dp = w, \tag{23}$$

where $x_p^* = R^{-1}(R_p^*)$ is the inverse function of $R_p^* = R(x_p^*)$. Since we have assumed $R'(x_p) < 0$, and R_p^* is increasing in R_0^* by Proposition 3, we have $F'(R_0^*) < 0$, which implies

$$\frac{dR_0^*}{dw} = \frac{1}{F'(R_0^*)} < 0.$$

Hence, an increase in the aggregate supply of savings w leads to a decrease in the safe rate R_0^* and consequently in the rates R_p^* charged to entrepreneurs of all types p. This, in turn, implies a higher investment x_p^* for all types p.

Since the marginal type p^* in (22) is decreasing in the equilibrium safe rate R_0^*, the originate-to-distribute region will be larger. Moreover, by Proposition 3 the increase in w will reduce the monitoring intensity m_p^* of originate-to-hold banks, so they will be riskier.

We can summarize these results as follows.

Proposition 4 *An increase in the aggregate supply of savings w leads to*

1. *A reduction in the safe rate R_0^* and in the loan rates R_p^* of all types of entrepreneurs.*

2. *An increase in investment x_p^* and hence in the size of banks' lending to all types of entrepreneurs.*

3. *An expansion of the range* $[0, p^*]$ *of entrepreneurs that borrow from banks that originate-to-distribute, and a shrinkage of the range* $[p^*, 1]$ *of entrepreneurs that borrow from banks that originate-to-hold.*

4. *A reduction in interest rate spreads* $R_p^* - R_0^*$.

5. *An reduction in the monitoring intensity* m_p^* *(and hence an increase in the probability of failure* $p - m_p^*$*) of originate-to-hold banks.*

We can illustrate these results for the case where the monitoring cost function is given by (2) and the relationship between the success return R_p and the aggregate investment of entrepreneurs of type p is given by (13).

When the monitoring cost function is quadratic, solving the first-order condition (19) we obtain the following equilibrium monitoring intensity of originate-to-hold banks

$$m_p^* = p - \left(1 - \sqrt{\frac{R_0^*}{\gamma}}\right) = p - p^*, \text{ for } p > p^*.$$

This implies $p - m_p^* = p^*$, so *all banks that originate-to-hold have the same probability of failure*, which equals the type p^* of the marginal entrepreneur. Thus, in this case p^* fully characterizes the risk of the banking system.

Substituting this result in (18) gives the following equilibrium loan rates for types $p > p^*$

$$R_p^* = \gamma(p - p^*) + \frac{R_0^*}{1 - p^*}. \tag{24}$$

Thus, with the quadratic monitoring cost function, equilibrium loan rates R_p^* (and spreads $R_p^* - R_0^*$) for originate-to-hold banks are linear in the risk type p.

Figure 2 shows the effects of an increase in the aggregate supply of savings w. Equilibrium variables before the change are indicated with a star and represented by solid lines, while equilibrium variables after the change are indicated with two stars and represented by dashed lines. The horizontal axis of the four panels represents the entrepreneurs' types p. They all show the shift in the position of the marginal type from p^* to p^{**}. The intuition for this shift is straightforward. The reduction in interest rate spreads associated with the increase in w

implies that banks lending to entrepreneurs of types slightly above p^* will have an incentive to reduce their monitoring. But since m_p^* is close to zero they will move to a corner solution with $m_p^{**} = 0$, so the originate-to-distribute region will expand.

Panel A shows the effect on equilibrium loan rates. The increase in w shifts downwards the function R_p^* to R_p^{**}. The intercept of these functions is the interest rate charged to entrepreneurs of type $p = 0$ (the safe rate), which goes down from R_0^* to R_0^{**}. To the left of the marginal types p^* and p^{**}, loan rates are convex in p (and given by (17)), while to the right of these points they are linear (and given by (24)).

Panel B shows the effect on equilibrium investment allocations. The increase in w shifts upwards the function x_p^* to x_p^{**}. The total amount of lending by banks that originate-to-distribute is clearly increasing, since banks that were initially using this mode of finance will increase their lending, and some banks that were monitoring their borrowers are now originating-to-distribute. The effect on the total amount of lending by banks that originate-to-hold is in principle ambiguous, because fewer banks monitor their borrowers although they become bigger. In our parameterization, lending by banks that originate-to-hold is also increasing, but the proportion of total lending that is accounted for by them goes down. In other words, these results show that a savings glut increases the relative size of the originate-to-distribute (shadow) banking system.

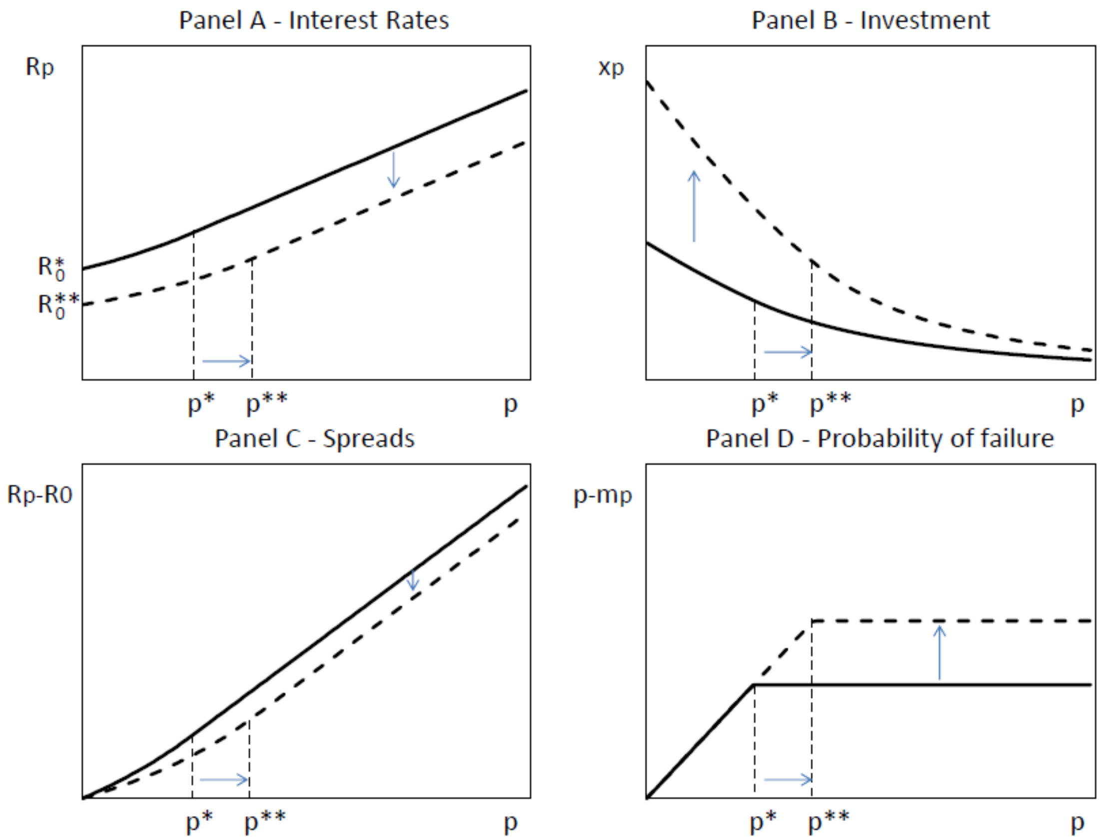

Figure 2. Effects of an increase in the supply of savings

This figure shows the effects of an increase in the supply of savings on equilibrium loan rates (Panel A), investment (Panel B), spreads (Panel C), and the probability of failure (Panel D) for different types of entrepreneurs. Solid (dashed) lines represent equilibrium values before (after) the increase in savings.

Panel C shows the effects on equilibrium spreads. As stated in Proposition 4, interest rate spreads go down from $R_p^* - R_0^*$ to $R_p^{**} - R_0^{**}$. Since equilibrium loan rates for originate-to-hold banks are linear in p with a slope equal to γ (see (24)), it follows that for types riskier than p^{**} spreads will be reduced by a constant amount.

Finally, Panel D shows the effect on equilibrium probabilities of bank failure. The shift of entrepreneurs with types in the interval between p^* and p^{**} from monitoring to nonmonitoring banks means that their probability of default will go up. Also, banks that originate-to-hold will increase their probability of failure from $p - m_p^* = p^*$ to $p - m_p^{**} = p^{**} > p^*$. Thus, the increase in the aggregate supply of savings w has an *extensive margin effect* due to the shift of originate-to-distribute banks toward riskier entrepreneurs (shown by the horizontal arrow), and an *intensive margin effect* due to the reduction in the intensity of monitoring by originate-to-hold banks (shown by the vertical arrow). Hence, we conclude that *a savings glut increases the risk of the banking system.*

We have so far analyzed the effects of an exogenous shock to the supply of savings w. However, the same effects obtain when there is an exogenous decrease in the demand for investment, which in the context of our model may be captured by a decrease of parameter a of the function R_p in (13). Substituting this function into the market clearing condition (21) gives

$$\int_0^1 x_p^* \, dp = a \int_0^1 \left(R_p^*\right)^{-\sigma} \, dp = w.$$

Clearly, equilibrium allocations depend on the ratio w/a, so we conclude that the effects of an increase in the aggregate supply of savings are identical to the effects of fall in the demand for investment.

Summing up, we have embedded a partial equilibrium model of bank finance into a simple general equilibrium model of the determination of equilibrium interest rates. The results show that an increase in the supply of savings (or a fall in the demand for investment) reduces interest rates and interest rate spreads, increases the relative size of the originate-to-distribute (shadow) banking system, and increases the probability of failure of the originate-to-hold (traditional) banks. These results provide a consistent explanation of a number of stylized facts of the period preceding the 2007-2009 financial crisis; see, for example, Brunnermeier (2009). They also provide an explanation of the way in which changes leading to a reduction in the equilibrium real rate of interest, as those noted by Summers (2014), can be linked to an increase in financial instability.

3.2 Efficiency of equilibrium

We next address whether the equilibrium of the model is (constrained) efficient, that is whether a social planner subject to the same moral hazard problem as the banks could improve upon the equilibrium allocation. We show that the equilibrium is constrained inefficient: the social planner would shift investments toward safer entrepreneurs.

To characterize the second-best efficient allocation we first have to derive the objective function of the social planner. This requires computing the social surplus S_p associated with output x_p of entrepreneurs of type p (which obtains with probability $1-p+m_p$). To do this we can use our previous derivation of the function $R(x_p)$ in (13) from either the demand of a set of final good producers that use the entrepreneurs' output as an intermediate input, or from the demand of a representative consumer.

In the case where $R(x_p)$ is derived from the demand of a set of final good producers we have

$$S_p = \int_{R_p}^{\infty} (\theta_p - R_p)\, g(\theta_p) d\theta_p + (R_p - B_p)\, x_p + B_p x_p,$$

where the first term are the profits of the final good producers, the second term the profits of the banks, and the third are the revenues of the investors. Then using (12) and the assumption $g(\theta_p) = a\sigma\,(\theta_p)^{-(\sigma+1)}$ this simplifies to

$$S_p = \frac{a\sigma}{\sigma-1}\left(\frac{x_p}{a}\right)^{\frac{\sigma-1}{\sigma}} = \frac{\sigma}{\sigma-1} R(x_p) x_p. \tag{25}$$

In the case where $R(x_p)$ is derived from the demand of a representative consumer we have

$$S_p = \left[\frac{a\sigma}{\sigma-1}\left(\frac{x_p}{a}\right)^{\frac{\sigma-1}{\sigma}} - R_p x_p\right] + (R_p - B_p)\, x_p + B_p x_p,$$

where the first term is the surplus of the representative consumer, the second term the profits of the banks, and the third are the revenues of the investors, which also gives (25).

A *constrained efficient allocation* $\{\widehat{x}_p\}_{p\in[0,1]}$ is an allocation that maximizes expected social surplus net of monitoring costs

$$\int_0^1 \left[(1-p+m_p)S_p - c(m_p)x_p\right]\, dp,$$

subject to the condition that characterizes optimal monitoring by the social planner

$$m_p = \max\left\{ m \in [0,p] \mid c'(m) + \frac{\widehat{R}_0}{1-p+m} \leq \widehat{R}_p \right\}, \text{ for all } p \in [0,1], \tag{26}$$

and the market clearing constraint

$$\int_0^1 \widehat{x}_p \, dp = w. \tag{27}$$

Since the social planner is subject to the same informational constraints as the banks, condition (26) follows from the characterization of the optimal contract between the bank and the investors in Proposition 1. Condition (27) ensures that the aggregate demand for investment is equal to the aggregate supply of savings w.

For those types p for which $m_p = 0$ (those sufficiently close to $p = 0$), differentiating the Lagrangian with respect to x_p gives

$$(1-p)R_p = \lambda, \tag{28}$$

where λ is the Lagrange multiplier associated with the market clearing constraint (27), and we have used the expression for R_p in (13). For $p = 0$ this implies $R_0 = \lambda$, so the Lagrange multiplier is the safe interest rate. Hence, for small values of p the expected return $(1-p)R_p$ equals the funding cost R_0.

For those types p for which $m_p > 0$, differentiating the Lagrangian with respect to x_p gives

$$(1-p+m_p)R_p - c(m_p) + \left[\frac{\sigma}{\sigma-1}R_p - c'(m_p)\right] x_p \frac{dm_p}{dR_p}\frac{dR_p}{dx_p} = \lambda, \tag{29}$$

where we have the expressions for R_p and S_p in (13) and (25). In contrast with the equilibrium allocation, the constrained efficient allocation is now characterized by a loan rate R_p that is strictly greater than the lowest feasible rate $\underline{R}_p$ defined in (15). To see this, notice that if the monitoring intensity m_p satisfied the first-order condition (19) that characterizes the lowest feasible rate, then we would have $dm_p/dR_p = \infty$, in which case (29) could not hold.[9] Hence, we conclude that compared to the equilibrium allocation, the constrained efficient allocation

[9] Notice that by (26) the term in square brackets in (29) is positive.

is characterized by lower investment and higher loan rates for riskier entrepreneurs, which by the market-clearing condition (27) implies higher investments and lower rates for safer entrepreneurs.

Figure 3 illustrates the constrained efficient allocation together with the corresponding equilibrium allocation for the case in which the monitoring cost function is quadratic (and, as before, we assume $a = 1$ and $\gamma =?$). The constrained efficient allocation is such that there exists $\widehat{p}$ such that $(1-p)\widehat{R}_p = \widehat{R}_0$ for $p \leq \widehat{p}$, and $\widehat{R}_p > \underline{R}_p$ for $p > \widehat{p}$. Panel A shows that the social planner would reallocate investments toward safer entrepreneurs, which reduces the safe interest rate and increases the loan rate charged to riskier entrepreneurs. The associated increase in spreads implies an increase in the monitoring of these entrepreneurs, whose projects will become safer, a result shown in Panel B.

TBC

4 Endogenous Booms and Busts

We have so far analyzed the equilibrium of a static model for a given aggregate supply of savings and shown how an exogenous change in this supply affects the risk of the banking system. This section analyzes a dynamic extension of the static model in which investors are infinitely lived and the aggregate supply of savings is endogenous. Specifically, the supply of savings at any date is the outcome of agents' decisions at the previous date together with the realization of a systematic risk factor that affects the return of entrepreneurs' projects.

The dynamic model generates endogenous booms and busts. The intuition is straightforward: the accumulation of savings leads to a reduction in interest rates and interest rate spreads and an increase in risk-taking that eventually materializes in a bust, which reduces savings, increasing interest rates and interest rate spreads and reducing risk-taking, starting again the process of wealth accumulation that leads to a boom.

At each date t there is a continuum of one-period-lived entrepreneurs of types $p \in [0, 1]$ that have investment projects that can only be funded by banks. As before, we assume that banking sector is contestable and that there is a single bank that lends to entrepreneurs of type p at date t, choosing the monitoring intensity $m_{pt} \in [0, p]$. The project of an entrepreneur of type p yields at date t a return $R_{pt} = R(x_{pt})$ with probability p and zero with probability $1 - p + m_{pt}$, where x_{pt} denotes the aggregate investment of entrepreneurs of type p at date t, and $R(x)$ is given by (13).

At each date t there is a continuum of measure w_t of infinitely-lived risk-neutral atomistic investors with unit wealth. Investors have a discount factor $\beta \in (0, 1)$ and the period utility function is given by $u(c_t) = c_t$. These investors fund the banks which in turn fund the entrepreneurs' projects. To simplify the presentation, we assume that banks are run by penniless one-period-lived bankers that consume the profits that they may obtain before they die. Thus, banks have no equity capital and bank profits do not contribute to the accumulation of wealth.

To describe the dynamics of wealth accumulation we need a model of the realization of project returns. We will maintain the assumption that the returns of the projects of entrepreneurs of each type p are perfectly correlated, but will assume that project returns are correlated across types. Specifically, we will use the *single risk factor model* of Vasicek (2002) in which the outcome of the projects of entrepreneurs of type p is driven by the realization of a latent random variable

$$y_{pt} = -\Phi^{-1}(p - m_{pt}) + \sqrt{\rho}\, z_t + \sqrt{1 - \rho}\, \varepsilon_{pt}, \tag{30}$$

where z_t is a *systematic risk factor* that affects all types of entrepreneurs, ε_{pt} is an idiosyncratic risk factor that only affects the projects of entrepreneurs of type p, $\rho \in (0, 1)$ is a parameter that determines the extent of correlation in the returns of the projects of entrepreneurs of different types, $\Phi(\cdot)$ denotes the cdf of a standard normal random variable, and $\Phi^{-1}(\cdot)$ its inverse. It is assumed that z_t and ε_{pt} are standard normal random variables, independently distributed from each other as well as, in the case of ε_{pt}, across types. They are also independent over time. The projects of entrepreneurs of type p fail at date t when $y_{pt} < 0$. Hence, their probability of failure is

$$\Pr(y_{pt} < 0) = \Pr\left[\sqrt{\rho}\, z_t + \sqrt{1 - \rho}\, \varepsilon_{pt} < \Phi^{-1}(p - m_{pt})\right] = p - m_{pt}.$$

The dynamic behavior of aggregate wealth is then given by

$$w_{t+1} = G(w_t, z_t) = \int_0^1 \delta(p - m_{pt}, z_t) x_{pt} B_{pt} \, dp, \tag{31}$$

where

$$\delta(p - m_{pt}, z_t) = \Pr(y_{pt} \geq 0 \mid z_t) = \Phi\left(\frac{\sqrt{\rho}\ z_t - \Phi^{-1}(p - m_{pt})}{\sqrt{1-\rho}}\right). \tag{32}$$

The integrand in the right-hand side of (31) is the conditional (on the realization of the systematic risk factor z_t) probability of success of the projects of entrepreneurs of type p at date t multiplied by the payment to investors in case of success, which is equal to the product of the investment x_{pt} by the interest rate at which they lend to the corresponding bank, B_{pt}. Since the systematic risk factor z_t is a random variable, the dynamic behavior of aggregate wealth will also be random.

For expositional purposes we assume that investors can either consume their unit wealth, invest it in the bank lending to entrepreneurs of type $p = 0$, or invest it in the bank lending to entrepreneurs of an arbitrary type $p > 0$.[10] Let $s_0 \in [0, 1]$ and $s_p \in [0, 1]$ denote the amounts invested in the two banks, and $c = 1 - s_0 - s_p \in [0, 1]$ the amount consumed. The Bellman equation is then given by

$$v(w_t) = \max_{(s_0, s_p)} \left\{1 - s_0 - s_p + \beta \left[s_0 R_{0t} E\left[v(w_{t+1})\right] + s_p B_{pt} E\left[\delta(p - m_{pt}, z_t) v(w_{t+1})\right]\right]\right\}. \tag{33}$$

Since $\lim_{x \to 0} R(x) = \infty$, in equilibrium we must have $s_0 > 0$ and $s_p > 0$. Then, differentiating the right-hand side of (33) with respect to s_0 and s_1, equating to zero the resulting expressions, and subtracting one from the other, gives the following condition

$$R_{0t} E\left[v(w_{t+1})\right] = B_{pt} E\left[\delta(p - m_{pt}, z_t) v(w_{t+1})\right]. \tag{34}$$

This condition states that investors must be indifferent between lending to the two banks.

In equilibrium it must be the case that $v(w) \geq 1$, since the investor can always set $s_0 = s_p = 0$, consuming all her wealth, which gives $u(1) = 1$. It must also be the case that $v(w) > 1$ only if $c = 1 - s_0 - s_p = 0$, since if $c > 0$ the marginal utility of lending to any

[10] As will be clear below, restricting investment to only two banks is without loss of generality, as investors will be indifferent among any of the banks.

of the two banks must be equal to the marginal utility of consumption which is one. Let us now define

$$\widehat{w} = \inf \{w \mid v(w) = 1\}. \tag{35}$$

Clearly, we have $v(w) = 1$ for all $w \geq \widehat{w}$. Thus, when $w < \widehat{w}$ the value of one unit of wealth is greater than one and investors do not consume, while when $w \geq \widehat{w}$ the value of one unit of wealth is equal to one and they invest $\widehat{w}$ and devote the difference $w - \widehat{w}$ to consumption. Hence, the aggregate consumption of investors is given by

$$c(w_t) = \begin{cases} w_t - \widehat{w}, & \text{for } w_t \geq \widehat{w}, \\ 0, & \text{for } w_t < \widehat{w}. \end{cases} \tag{36}$$

Substituting the indifference condition (34) into the Bellman equation (33) gives

$$v(w_t) = \beta R_{0t} E\left[v(w_{t+1})\right] = \beta B_{pt} E\left[\delta(p - m_{pt}, z_t) v(w_{t+1})\right], \tag{37}$$

which implies the fundamental pricing equation

$$E\left[\frac{\beta v(w_{t+1})}{v(w_t)} \delta(p - m_{pt}, z_t) B_{pt}\right] = 1, \tag{38}$$

where $\beta v(w_{t+1})/v(w_t)$ is the stochastic discount factor. Given that the expected value of the stochastic discount factor equals the inverse of safe rate R_{0t}, and that $E\left[\delta(p - m_{pt}, z_t)\right] = \Pr(y_{pt} \geq 0) = 1 - p + m_{pt}$, we have

$$E\left[\frac{\beta v(w_{t+1})}{v(w_t)} \delta(p - m_{pt}, z_t) B_{pt}\right] = \frac{(1 - p + m_{pt}) B_{pt}}{R_{0t}} + Cov\left[\frac{\beta v(w_{t+1})}{v(w_t)}, \delta(p - m_{pt}, z_t) B_{pt}\right],$$

so the pricing equation (38) implies

$$(1 - p + m_{pt}) B_{pt} - R_{0t} = -R_{0t} Cov\left[\frac{\beta v(w_{t+1})}{v(w_t)}, \delta(p - m_{pt}, z_t) B_{pt}\right].$$

Now since $w_{t+1} = G(w_t, z_t)$ and $\delta(p - m_{pt}, z_t)$ are both increasing in the systematic risk factor z_t, and $v'(w_{t+1}) \leq 0$, with strict inequality for low values of w_{t+1}, we conclude that the covariance term is negative, which implies

$$(1 - p + m_{pt}) B_{pt} - R_{0t} > 0,$$

for all $p > 0$. Thus, *investors require positive risk premia for funding the risky banks.* In other words, they behave as if they were risk-averse.

Following the same steps as in the analysis of the static model in Section 3, and solving for B_{pt} in (34), one can show that banks lending to entrepreneurs of types $p > 0$ will set the lowest feasible loan rate, which is given by

$$\underline{R}_{pt} = \min_{m_{pt} \in [0,p]} \left(c'(m_{pt}) + \frac{R_{0t} E\left[v(w_{t+1})\right]}{E\left[\delta(p - m_{pt}, z_t) v(w_{t+1})\right]} \right). \tag{39}$$

Notice that in the static model we have $v(w_{t+1}) = 1$, which implies $E\left[\delta(p - m_{pt}, z_t) v(w_{t+1})\right] = 1 - p + m_{pt}$, so (39) becomes (15).

We are now ready to define an equilibrium, which requires to specify the investment x_{pt} of the different types of entrepreneurs, and hence the rates $R_{pt} = R(x_{pt})$ at which they will borrow from the banks, the rates B_{pt} at which the banks will borrow from the investors, and their monitoring intensity m_{pt}. All these variables depend on the wealth w_t of the investors, which is the state variable of the dynamic model. The equilibrium also requires to specify the value function of the investors $v(w_t)$, their aggregate consumption decision $c(w_t)$, and the dynamics of wealth accumulation.

Formally, an *equilibrium* is an array $\left\{x_p^*(w_t), R_p^*(w_t), B_p^*(w_t), m_p^*(w_t)\right\}_{p \in [0,1]}, v(w_t), c(w_t), G(w_t, z_t)$ such that

1. Entrepreneurs' investment decisions satisfy $R_p^*(w_t) = R\left(x_p^*(w_t)\right)$,

2. Banks' lending rates $R_p^*(w_t)$ equal the lowest feasible rates $\underline{R}_{pt}$ in (39),

3. Banks' borrowing rates $B_p^*(w_t)$ satisfy the fundamental pricing equation (38),

4. Banks' monitoring intensity $m_p^*(w_t)$ solves (39),

5. The value function $v(w_t)$ satisfies (37),

6. The consumption function $c(w_t)$ satisfies (36),

7. The investors' wealth w_t evolves according to (31), and

8. The market clears

$$\int_0^1 x_p^*(w_t)\ dp = w_t - c(w_t).$$

We can illustrate the equilibrium of the dynamic model using the parameterization in Section 3.[11] An interesting result that also obtains here with the quadratic monitoring cost function (2) is that there exists a marginal type p_t^* such that banks lending to entrepreneurs of types $p \leq p_t^*$ will originate-to-distribute, setting $m_{pt}^* = 0$, and banks lending to types $p > p_t^*$ will originate-to-hold, setting $m_{pt}^* = p - p_t^*$. This implies $p - m_p^* = p^*$, so *all banks that originate-to-hold have the same probability of failure*, which equals the type p_t^* of the marginal entrepreneur.[12]

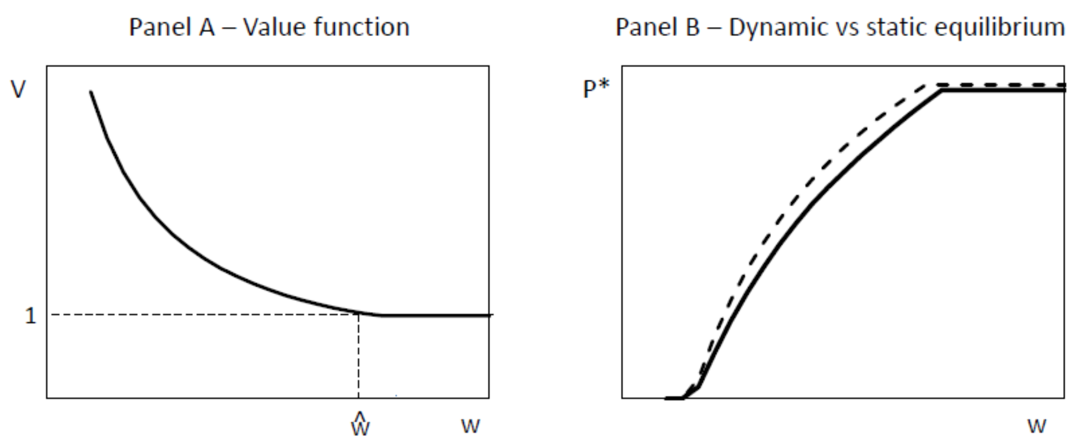

Figure 6. Value function and dynamic vs static equilibrium

This figure shows the value of one unit of wealth in the dynamic model (Panel A), and compares the dynamic (solid) with static (dashed) probability of failure of the traditional banks (Panel B) for a range of values of the state variable w.

[11] We assume a discount factor $\beta = 0.96$ and a correlation parameter $\rho = 0.15$.

[12] To prove this result, suppose that the solution to (39) for some p is such that $m_{pt}^* > 0$. Then, m_{pt}^* satisfies the first-order condition

$$\gamma + \frac{d}{dm_{pt}} \left(\frac{R_{0t}^* E\left[v(w_{t+1})\right]}{E\left[\delta(p - m_{pt}, z_t) v(w_{t+1})\right]} \right) \Bigg|_{m_{pt} = p - p_t^*} = 0.$$

But given the functional form of $\delta(p - m_{pt}, z_t)$ in (32), it follows that this condition is also satisfied for any $p \geq p_t^*$ (so that the corresponding $m_{pt}^* = p - p_t^* \geq 0$). Moreover, for the same reason, there cannot be an interior solution for $p < p_t^*$, which proves the result.

The indifference condition (34) and the condition (39) that characterizes equilibrium loan rates imply

$$R^*_{pt} = \gamma m^*_{pt} + B^*_{pt}.$$

Hence, banks' lending and borrowing rates, R^*_{pt} and B^*_{pt}, coincide for $p \leq p^*_t$, that is for banks that originate-to-distribute, and satisfy $R^*_{pt} > B^*_{pt}$ for $p > p^*_t$, that is for banks that originate-to-hold. In this case, as discussed in Section 3, the intermediation margin $R^*_{pt} - B^*_{pt}$ covers the monitoring cost and leaves some positive profits for the banks.

Panel A of Figure 6 shows the value function $v(w)$, which is decreasing and convex for $w < \widehat{w}$, and satisfies $v(w) = 1$ for $w \geq \widehat{w}$. As noted above, the fact that $v'(w) \leq 0$, with strict inequality for low values of w, implies the existence of positive risk premia, like in the static model with risk-averse investors analyzed in Section 4.2. The positive risk premia opens up interest rate spreads, increasing the value of monitoring and consequently the comparative advantage of originate-to-hold banks.

Panel B of Figure 6 illustrates this effect by comparing, for different values of the state variable w, the marginal type p^* for the dynamic model (solid line) with that of the static model with risk-neutral investors (dashed line). The former is everywhere below the latter, except for low values of w when $p^* = 0$ in both models. This means that the dynamic model features a smaller relative size of the originate-to-distribute banking system, and a lower probability of failure of the originate-to-hold banks. In other words, *the forward-looking behavior of investors contributes to the stability of the banking system.*

The dynamic model yields a number of potentially testable relationships between aggregate variables. In order to highlight some of these relationships, in what follows we consider a sample realization of the (iid) systematic risk factor z_t and look at the corresponding evolution of investors' wealth w_t over time together some interesting variables. The black line of Panels A-D of Figure 7 represents w_t, and is measured in the left-hand-side axis, and the horizontal dashed line shows the value $\widehat{w}$ above which investors consume $w_t - \widehat{w}$.

In Panel A the dark grey line plots the total amount of lending by (traditional) banks that originate-to-hold, that is $\int_{p^*_t}^1 x^*_p(w_t)\ dp$, and the light grey line plots the total amount of lending by (shadow) banks that originate-to-distribute, that is $\int_0^{p_t^*} x_p^*(w_t)\ dp$. Although lending by both traditional and shadow banks is positively correlated with investors' wealth, most of the variation in w_t is channeled through shadow banks. In other words, *the shadow banking system is highly procyclical*, a result that is consistent with the evidence in Pozsar, Adrian, Ashcraft, and Boesky (2012) that shows that shadow bank liabilities grew much faster than traditional banking liabilities in the run-up to the crisis, and contracted substantially following the 2007 peak.

The grey lines in Panel B show the evolution of the risk premia $(1 - p + m_{pt}^*)B_{pt}^* - R_{0t}^*$, measured in percentage points in the right-hand axis, for two types $p = 0.05$ (light grey) and $p = 0.1$ (dark grey) that are always funded by shadow banks. Risk premia are negatively correlated with w_t, and are higher for the riskier type. The fact that risk premia go down when wealth accumulates means that the dynamic model yields *countercyclical risk premia*. Thus, investors behave as if they were less risk-averse (or have greater risk appetite) during financial booms, although their underlying (risk-neutral) preferences do not change.

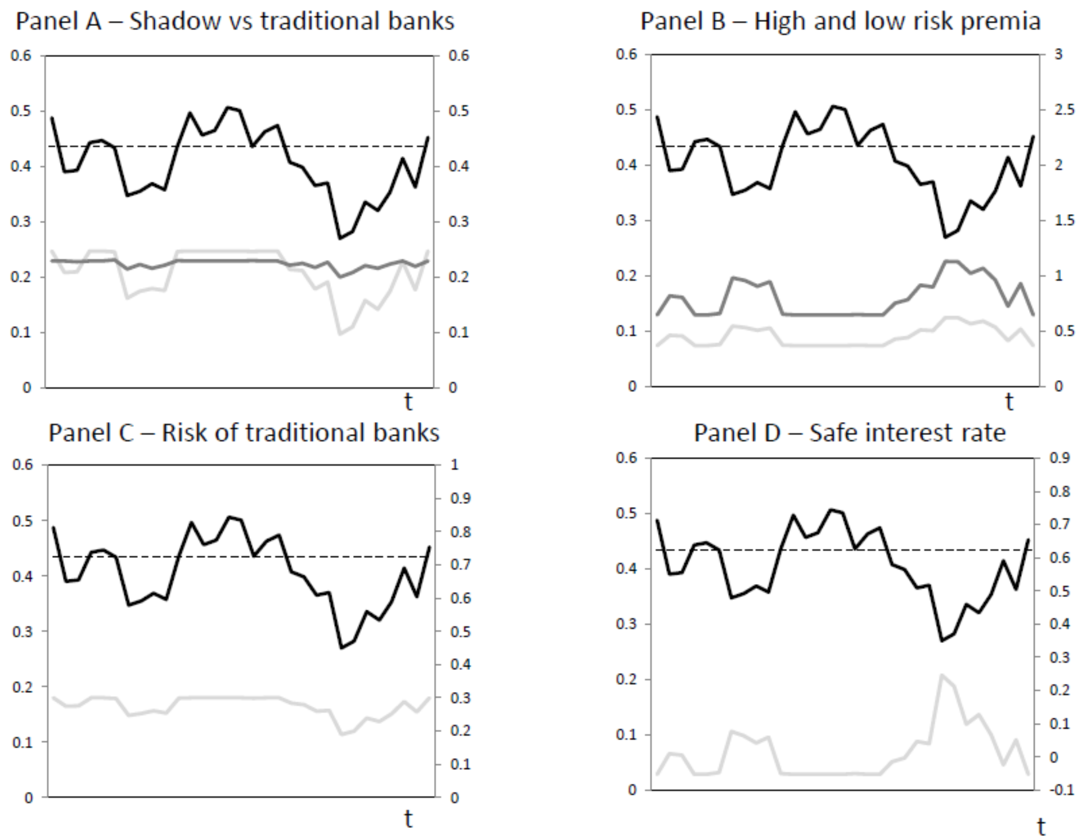

Figure 7. Model dynamics

The black line in all panels shows the dynamic evolution of investors' wealth for a sample realization of the systematic risk factor, and the dashed line shows the threshold level of wealth above which investors consume. Panel A plots the total amount of lending by traditional banks (dark grey) and shadow banks (light grey), Panel B the risk premia for investors in two different shadow banks, Panel C the probability of failure of the traditional banks, and Panel D the safe rate (in net terms).

The grey line in Panel C shows the evolution of the type p_t^* of the marginal entrepreneur, measured in the right-hand axis. This variable is positively correlated with w_t. Thus, in line with the results of the static model, higher wealth increases the relative size of the shadow banking system, and increases the probability of failure of the traditional banks.

Finally, the grey line in Panel D shows the evolution of the safe rate R_{0t}^*, measured in net terms in the right-hand axis. This variable is negatively correlated with w_t. Moreover, *the safe rate can be below the discount rate $1/\beta$ and even become negative in booms.*[13] The intuition is that the expectation of positive returns in the future (when the economy is hit by a negative shock and wealth is very valuable) makes investors willing to forgo current consumption, which lowers the safe rate.

As previously explained, the rationale for the observed dynamics is as follows: good realizations of the systematic risk factor result in higher aggregate wealth which reduces interest rates and interest rate spreads, increases the size of the shadow banking system, and increases the probability of failure of the traditional banks, making the banking system especially vulnerable to a bad realization of the systematic risk factor. Therefore, our model provides a framework for understanding the emergence of endogenous boom and bust cycles, as well as the procyclical nature of the shadow banking system, the existence of countercyclical risk premia, and the low levels of interest rates and interest rate spreads during booms.

5 Conclusion

This paper presents a general equilibrium model of the connection between real interest rates, credit spreads, and the structure and the risk of the banking system. Banks intermediate between a heterogeneous set of entrepreneurs and a set of investors characterized by a fixed aggregate supply of savings. We assume that all agents are risk-neutral and that banks can monitor entrepreneurs' projects at a cost, but this is not observed by investors. This moral hazard problem is the key friction that drives the results of the model. We also assume that project returns are decreasing in the aggregate investment of entrepreneurs of each type, and that the market for lending to entrepreneurs is contestable.

We first characterize the equilibrium of the model, showing that safer entrepreneurs will be funded by banks that do not monitor their projects and riskier entrepreneurs by banks that monitor them. We assume that monitoring requires keeping the loans in the banks' books, and for this reason we associate nonmonitoring banks with intermediaries that originate-to-distribute (called shadow banks) and monitoring banks with intermediaries that

[13] Notice that for $w_t \geq \widehat{w}$ we have $v(w_t) = 1$ and $E\left[v(w_{t+1})\right] > 1$, so (37) implies $\beta R_{0t} < 1$, i.e. $R_{0t} < 1/\beta$.

originate-to-hold (called traditional banks). We then analyze the effects of an increase in the supply of savings, showing that it will lead to a reduction in interest rates and interest rate spreads, an expansion of the relative size of the shadow banking system, and a reduction in the monitoring intensity and hence an increase in the probability of failure of the traditional banks.

We extend our static model to a dynamic setup in which the aggregate supply of savings at any date is the outcome of agents' decisions at the previous date together with the realization of a systematic risk factor that affects the return of entrepreneurs' projects. The dynamic model generates endogenous booms and busts: the accumulation of savings leads to a reduction in interest rates and interest rate spreads and an increase in risk-taking that eventually materializes in a bust, which reduces savings, starting again the process of wealth accumulation that leads to a boom. The model also yields a number of empirically relevant results such as the procyclical nature of the shadow banking system, the existence of countercyclical risk premia, and the low levels of interest rates and interest rate spreads leading to the buildup of risks during booms.

These results provide a theoretical explanation for the relationship noted by Rajan (2005) between high savings, low real interest rates, and the incentives of financial intermediaries to search for yield. They also rationalize the way in which changes leading to a reduction in real interest rates, as those noted by Summers (2014), can be associated with an increase in financial instability. Moreover, the results provide a rationale for a number of empirical facts in the run-up of the 2007-2009 financial crisis.

We would like to conclude with a number of remarks. First, our model builds on the assumption that banks monitor their borrowers, an idea that traces back to the seminal work of Diamond (1984).[14] Still, a similar story could be constructed if banks increase the quality its pool of loan applicants by screening them.[15] Second, the paper entirely focuses on debt

[14] The relevance of this assumption has been recently documented by Wang and Xia (2014), who show that "banks active in securitization impose looser covenants on borrowers at origination" and "after origination, these borrowers take on substantially more risk than do borrowers of non-securitization-active banks."

[15] This setup would be similar to the one in Helpman, Itskhoki and Redding (2010), where firms can increase the average ability of the workers they hire by paying a screening cost.

finance, abstracting from the possibility of funding the banks with (inside) equity. Since equity finance would strengthen the banks' monitoring incentives, adding this possibility would require the introduction of a differential cost of equity.[16] Third, we only provide a positive analysis of the connection between real interest rates and financial stability, without analyzing the (constrained) efficiency of the equilibrium allocations, which would merit a separate paper. Finally, it is important to stress that we have constructed a model without any nominal frictions, so monetary policy is completely absent from our story of search for yield. Introducing nominal frictions would allow to study the connection between monetary policy and financial stability, a topic that would also merit a separate paper.

[16] See Martinez-Miera and Repullo (2016) for a model along these lines that is used to discuss the effects of different types of bank capital requirements.

Appendix

A Risk-averse investors

This Appendix extends the model of Section 3 to the case where investors are risk-averse instead of risk-neutral. This allows us to distinguish the effects of a change in the supply of savings from those of a change in investors' risk appetite.

Specifically, suppose that there is a continuum of measure w of atomistic investors with unit wealth that can be invested in only one bank (so we do not allow any portfolio diversification). Since each investor has a unit wealth, the measure of investors w is equal to the aggregate supply of savings.

Investors have a constant relative risk aversion utility function. Given that bank assets can yield a zero return, the coefficient of relative risk aversion is restricted to be between zero and one. Thus, we have

$$u(c) = c^{1/\alpha}, \tag{40}$$

where $\alpha \geq 1$. Risk-neutrality corresponds to the limit case $\alpha = 1$.

As in our baseline model, in equilibrium investors have to be indifferent between funding banks lending to different types of entrepreneurs. This implies that in the definition of an optimal contract between a bank lending to entrepreneurs of type p and the risk-averse investors, the participation constraint (6) becomes

$$\left(1 - p + m_p^*\right)\left(B_p^*\right)^{1/\alpha} = R_0^{1/\alpha},$$

which may be rewritten as

$$B_p^* = \frac{R_0}{(1 - p + m_p^*)^\alpha}. \tag{41}$$

Notice that the investors' expected payoff satisfies

$$(1 - p + m_p^*)B_p^* = \frac{R_0}{(1 - p + m_p^*)^{\alpha-1}} > R_0,$$

so they require positive risk premia.

Substituting the participation constraint (41) into the first-order condition (7) gives

$$c'(m_p) + \frac{R_0}{(1-p+m_p)^\alpha} = R. \tag{42}$$

As before, the function in left-hand side of (42) is convex in m_p. Let $\underline{R}_p$ denote the minimum value of this function. Then, we can follow the same steps as in the proof of Proposition 1 to show that bank finance is feasible if $R \geq \underline{R}_p$. In such case, the optimal contract between the bank and the investors is given by

$$m_p^* = \max\left\{m \in [0,p] \mid c'(m) + \frac{R_0}{(1-p+m)^\alpha} \leq R\right\}.$$

Thus, we have essentially the same characterization of the optimal contract as in the risk-neutral case analyzed in Section 2. The difference is that the convex function in the left-hand side of (42) is increasing in α, so risk aversion makes it more difficult to ensure the feasibility of bank finance.

Following the same steps as in Section 3, we can characterize the equilibrium of the model with risk-averse investors. In this equilibrium, the marginal type is given by

$$p^* = 1 - \left(\tfrac{\alpha R_0^*}{c''(0)}\right)^{\frac{1}{1+\alpha}}.$$

Notice that p^* is decreasing in the safe rate R_0^* (as before) and also in the risk-aversion parameter α. Thus, risk-aversion increases the value of monitoring and consequently the comparative advantage of originate-to-hold banks.

Figure 4 shows the effect of a reduction in the investors' risk aversion from $\alpha = 2$ to $\alpha = 1$ (risk-neutrality) for the same parameterization used in Section 3. As before, equilibrium variables before the change are indicated with a star and represented by solid lines, while equilibrium variables after the change are indicated with two stars and represented by dashed lines. The horizontal axis of the four panels represents the entrepreneurs' types p.

Panel A shows the effect on equilibrium loan rates. The reduction in risk aversion shifts the investors' preferences toward riskier assets, so loan rates go down for riskier entrepreneurs and increase for safer entrepreneurs. In particular, the safe rate will go up from R_0^* to R_0^{**}. The increase in the safe rate together with the reduction in loan rates for riskier entrepreneurs reduces the comparative advantage of originate-to-hold banks, which explains the shift in the position of the marginal type from p^* to p^{**}.

Panel B shows the effect on equilibrium investment allocations. The reduction in risk aversion produces a redistribution in the allocation of savings toward riskier entrepreneurs. Since the aggregate supply of savings is fixed, this means that investment in safer projects falls. However, the shift in the position of the marginal type from p^* to p^{**} implies that the effect on the relative size of the shadow banking system is ambiguous.

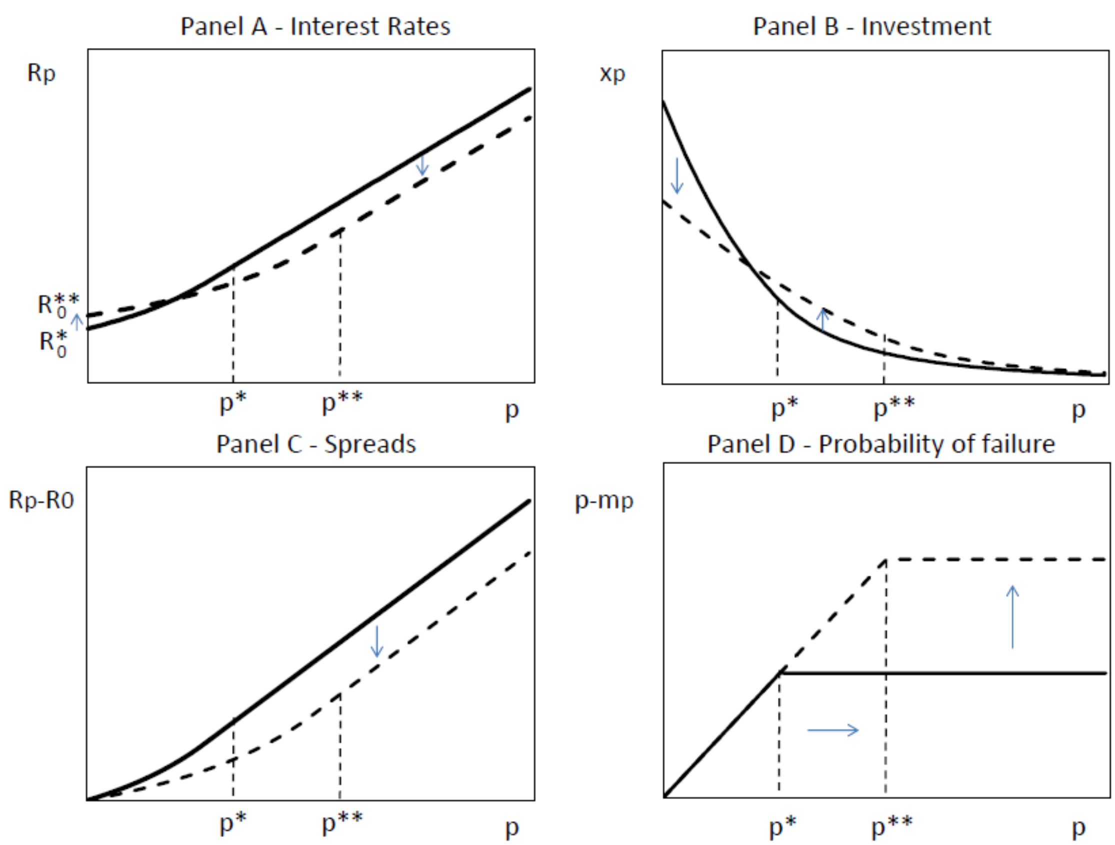

Figure 4. Effects of a reduction in investors' risk aversion

This figure shows the effects of a change in investors' preferences from risk aversion to risk neutrality on equilibrium loan rates (Panel A), investment (Panel B), spreads (Panel C), and the probability of failure (Panel D) for different types of entrepreneurs. Solid (dashed) lines represent equilibrium values before (after) the change in preferences.

Panel C shows the effects on equilibrium spreads. The results on loan rates imply that interest rate spreads go down from $R_p^* - R_0^*$ to $R_p^{**} - R_0^{**}$. This reduces the incentives to monitor and hence increases the probability of failure of originate-to-hold banks, which is shown in Panel D. As in the case of a savings glut, a reduction in risk-aversion has an extensive margin effect due to the shift of originate-to-distribute banks toward riskier borrowers (shown by the horizontal arrow), and an intensive margin effect due to the reduction in the intensity of monitoring by originate-to-hold banks (shown by the vertical arrow).

These results illustrate the differences between the effects of a savings glut from the effects of a reduction in the investors' risk appetite. Both changes lead to a reduction in interest rate spreads and an increase in the probability of failure of originate-to-hold banks, but there are some significant differences. A savings glut increases funding for all types of entrepreneurs and the size of the shadow banking system, while a fall in risk aversion reduces funding for safer entrepreneurs and has an ambiguous effect on the size of the shadow banking system. A simple way to empirically distinguish the two changes is to look at the effect on the equilibrium safe rate: it goes down in the case of a savings glut and it goes up in the case of a reduction in risk aversion.

B Proofs

Proof of Proposition 1 If $R < \underline{R}$, for any $m \in (0, p]$ we have

$$R - \frac{R_0}{1 - p + m} - c'(m) < 0,$$

which implies that the bank has an incentive to reduce monitoring m. But for $m = 0$ we have

$$R - \frac{R_0}{1 - p} - c'(0) < 0.$$

Using the assumption $c(0) = c'(0) = 0$, this implies $(1 - p)R - R_0 - c(0) < 0$, which violates the bank's participation constraint (5).

If $R \geq \underline{R}$, by the convexity of the function in the left-hand side of (8) there exist an interval $[m^-, m^*] \subset [0, p]$ such that

$$R - \frac{R_0}{1-p+m} - c'(m) \geq 0 \quad \text{if and only if} \quad m \in [m^-, m^*].$$

By our previous argument, for any $m \in (0, p]$ for which

$$R - \frac{R_0}{1-p+m} - c'(m) < 0,$$

the bank has an incentive to reduce monitoring m. Similarly, for any $m \in [0, p)$ for which

$$R - \frac{R_0}{1-p+m} - c'(m) > 0,$$

the bank has an incentive to increase monitoring m. Hence, there are three possible values of monitoring in the optimal contract: $m = m^*$, $m = m^-$, and $m = 0$ (when $m^- > 0$).

To prove that the bank prefers $m = m^*$, notice that our assumptions on the monitoring cost function together with the definition of m^* imply

$$\frac{d}{dm}\left[(1-p+m)R - c(m)\right] = R - c'(m) > R - c'(m^*) \geq B^* > 0,$$

for $m < m^*$. Hence, we have

$$\begin{aligned}(1-p+m^*)(R-B^*) - c(m^*) &= (1-p+m^*)R - R_0 - c(m^*) \\ &> (1-p+m)R - R_0 - c(m) = (1-p+m)(R-B) - c(m),\end{aligned}$$

for either $m = m^-$ or $m = 0$ (when $m^- > 0$).

Finally, to prove that the bank's participation constraint (5) is satisfied notice that

$$(1-p+m^*)R - R_0 - c(m^*) \geq (1-p+m^*)c'(m^*) - c(m^*) > (1-p)c'(m^*) > 0,$$

where we have used the the definition of m^* and the fact that $m^*c'(m^*) > c(m^*)$ by the convexity of the monitoring cost function. □

Proof of Proposition 2 Differentiating condition (8) for an interior level of monitoring $m^* \in (0, p)$ gives

$$\left(c''(m^*) - \frac{R_0}{(1-p+m^*)^2}\right) dm^* + \frac{1}{1-p+m^*} dR_0 - dR = 0.$$

By Proposition 1, when $R > \underline{R}$ and $m^* < p$ the term that multiplies dm^* is positive, which implies

$$\begin{aligned}
\frac{\partial m^*}{\partial R_0} &= -\frac{1}{1-p+m^*}\left(c''(m^*) - \frac{R_0}{(1-p+m^*)^2}\right)^{-1} < 0, \\
\frac{\partial m^*}{\partial R} &= \left(c''(m^*) - \frac{R_0}{(1-p+m^*)^2}\right)^{-1} > 0. \ \square
\end{aligned}$$

Proof of Proposition 3 To prove that higher types p will be characterized by higher spreads $R_p - R_0$ we simply differentiate (17) and apply the envelope theorem to (18). To prove that higher types $p > \widehat{p}$ will be characterized by higher monitoring m_p we differentiate the first-order condition (19) and use the assumption $c'''(m) \geq 0$.

To prove that increases in the safe rate R_0 lead to increases in the spreads $R_p - R_0$ for $p \leq \widehat{p}$ note that (17) implies

$$R_p - R_0 = \frac{pR_0}{1-p},$$

so spreads are linear in R_0. To prove that increases in the safe rate R_0 lead to increases in the spreads $R_p - R_0$ for $p > \widehat{p}$ we apply the envelope theorem to (18), which gives

$$\frac{d(R_p - R_0)}{dR_0} = \frac{1}{1-p+m_p} - 1 = \frac{p - m_p}{1-p+m_p} > 0.$$

Finally, to prove that monitoring m_p is increasing in the safe rate R_0 for $p > \widehat{p}$ we differentiate the first-order condition (19) and use the assumption $c'''(m) \geq 0$. $\square$

References

Adrian, T., P. Colla, and H. Shin (2013), "Which Financial Frictions? Parsing the Evidence from the Financial Crisis of 2007 to 2009," *NBER Macroeconomics Annual*, University of Chicago Press, 159-214.

Adrian, T., and H. Shin (2014), "Procyclical Leverage and Value-at-Risk," *Review of Financial Studies*, 27, 373-403.

Altunbas, Y., L. Gambacorta, and D. Marques-Ibanez (2014), "Does Monetary Policy Affect Bank Risk?," *International Journal of Central Banking*, 10, 95-135.

Bernanke, B. (2005), "The Global Saving Glut and the U.S. Current Account Deficit," Sandridge Lecture, Virginia Association of Economists, Richmond, Virginia.

Bernanke, B., and Mark Gertler (1979), "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review*, 79, 14-31.

Bernanke, B., M. Gertler, and S. Gilchrist (1996), "The Financial Accelerator and the Flight to Quality," *Review of Economics and Statistics*, 78, 1-15.

Boissay, F., F. Collard, and F. Smets (2015), "Booms and Banking Crises," *Journal of Political Economy, forthcoming.*

Brunnermeier, M. (2009), "Deciphering the Liquidity and Credit Crunch 2007-2009," *Journal of Economic Perspectives*, 23, 77-100.

Brunnermeier, M., T. Eisenbach, and Y. Sannikov (2012), "Macroeconomics with Financial Frictions: A Survey."

Brunnermeier, M., and Y. Sannikov (2014), "A Macroeconomic Model with a Financial Sector," *American Economic Review*, 104, 379-421.

Caballero, R., E. Farhi, and P.-O. Gourinchas (2008), "An Equilibrium Model of 'Global Imbalances' and Low Interest Rates," *American Economic Review*, 98, 358-393.

Dell'Ariccia, G., L. Laeven, and G. Suarez (2014), "Bank Leverage and Monetary Policy's Risk-Taking Channel: Evidence from the United States."

Diamond, D. (1984), "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51, 393-414.

Financial Stability Board (2014), *Global Shadow Banking Monitoring Report 2014.*

Gertler, M., and N. Kiyotaki (2010), "Financial Intermediation and Credit Policy in Business Cycle Analysis," in B. Friedman and M. Woodford (eds.), *Handbook of Monetary Economics*, Vol. 3, North Holland, 547-599.

Gilchrist, S., and E. Zakrajšek (2012), "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, 102, 1692-1720.

Hanson, S., A. Kashyap, and J. Stein (2011), "A Macroprudential Approach to Financial Regulation," *Journal of Economic Perspectives*, 25, 3–28.

He, Z., and A. Krishnamurthy (2012), "A Model of Capital and Crises," *Review of Economic Studies*, 79, 735-777.

Helpman, E., O. Itskhoki, and S. Redding (2010), "Inequality and Unemployment in a Global Economy," *Econometrica*, 78, 1239-1283.

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

Ioannidou, V., S. Ongena, and J.-L. Peydro (2015), "Monetary Policy, Risk-taking, and Pricing: Evidence from a Quasi-natural Experiment," *Review of Finance*, 19, 95-144.

Jimenez, G., S. Ongena, J.-L. Peydro, and J. Saurina (2014), "Hazardous Times for Monetary Policy: What Do Twenty-Three Million Bank Loans Say About the Effects of Monetary Policy on Credit Risk-Taking?," *Econometrica*, 82, 463–505.

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

Lopez-Salido, D., J. Stein, and E. Zakrajšek (2015), "Credit Market Sentiment and the Business Cycle."

Martinez-Miera, D., and R. Repullo (2016), "Comparing Flat and Risk-based Capital Requirements."

Melitz, M. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695-1725.

Muir, T. (2014), "Financial Crises and Risk Premia."

Pozsar, Z., T. Adrian, A. Ashcraft, and H. Boesky (2012), "Shadow Banking," Federal Reserve Bank of New York Staff Report No. 458.

Quadrini, V. (2011), "Financial Frictions in Macroeconomic Fluctuations," *Economic Quarterly*, Federal Reserve Bank of Richmond, 97, 209–254.

Rajan, R. (2005), "Has Financial Development Made the World Riskier?," Proceedings of the Jackson Hole Conference organized by the Federal Reserve Bank of Kansas City.

Repullo, R., and J. Suarez (2004), "Loan Pricing under Basel Capital Requirements," *Journal of Financial Intermediation,* 13, 496-521.

Repullo, R., and J. Suarez (2013), "The Procyclical Effects of Bank Capital Regulation," *Review of Financial Studies*, 26, 452-490.

Shin, H. S. (2012), "Global Banking Glut and Loan Risk Premium," *IMF Economic Review*, 60, 155-192.

Summers, L. (2014), "U.S. Economic Prospects: Secular Stagnation, Hysteresis, and the Zero Lower Bound," *Business Economics*, 49, 65-73.

Vasicek, O., (2002): "Loan Portfolio Value," *Risk*, 15, December, 160-162.

Wang, Y., and H. Xia (2014), "Do Lenders Still Monitor When They Can Securitize Loans?," *Review of Financial Studies*, 27, 2354-2391.