Department of Economics Discussion Paper Series

Productivity Booms, Bank Fragility, and Financial Crises

Artur Doshchyn

Number 1027
December, 2022

Department of Economics
Manor Road Building, Manor Road
Oxford, OX1 3UQ

ISSN 1471-0498

Productivity Booms, Bank Fragility, and Financial Crises*

Artur Doshchyn†

December 9, 2022

Abstract

While financial crises tend to be preceded by economic booms, most booms do not end with crises. Crises typically occur when booms are interrupted by persistent slowdowns in productivity growth. I develop a model in which risk of crisis emerges endogenously during boom because of increased fragility of the banking sector. Banks raise financing from households to invest in long-term projects, but their ability to do so is hindered by moral hazard, since bankers can misappropriate investors' funds. Demandable deposits create discipline by exposing misbehaving banks to runs, and thus help them increase external financing. Normally, banks finance themselves with a mix of equity and deposits that maximizes discipline, but ensures that they always remain solvent. But when growth prospects become sufficiently strong, worsening agency problems induce banks to intensify deposit financing, which enables a boom in credit, asset prices, and investment. If the anticipated growth fails to materialise, however, the excessive deposit financing leads to a systemic banking crisis.

*I am very grateful to Steve Bond, Chris Bowdler, Edouard Challe, Sarah Clifford, Martin Ellison, Alistair Macaulay, Benjamin Moll, Genevieve Nelson, and Oren Sussman for helpful discussions, comments and suggestions.

†University of Oxford and Linacre College, artur.doshchyn@economics.ox.ac.uk

I. Introduction

It is well-documented that major financial crises are typically preceded by booms in credit and asset prices[1] and are followed by protracted periods of slow growth.[2] Existing literature often studies booms as the result of financial liberalisation, excessive risk taking, or over-optimistic expectations.[3] Slow recoveries, on the other hand, are largely attributed to ensuing credit crunches and deleveraging during crises.[4] These mechanisms are undoubtedly important, but they cannot account for several important empirical observations that link financial cycles and the real economy. Specifically, Gorton and Ordoñez (2019), Dell'Ariccia et al. (2012) and Paul (2022) show that:

i. Starts of credit booms are often associated with accelerated productivity growth.

ii. Importantly, most credit booms do not end up in crises.

iii. When they do, however, financial crises tend to follow, rather than lead, persistent slowdowns in productivity growth, especially in its *trend component*.

In agreement with these observations, Cao and L'Huillier (2018) document that the three largest economic disasters in the developed world in the last 100 years – the Great Depression, the Japanese Slump of the 1990s, and the Great Recession – were all preceded by periods of major technological innovations and rapid productivity growth which ran out of steam *before* the respective crises. (In Appendix A, I provide a more detailed look at the 2008 Financial Crisis case). In sum, the evidence suggests that the recipe for a typical financial crisis includes a preceding period of strong economic growth interrupted by deteriorating macroeconomic fundamentals. But while weakening growth prospects may be unwelcome news, not least during booms, it is not clear why there have to be often devastating financial meltdowns in these circumstances. This leads to an important question which I study theoretically in this paper: How do booms make the economy susceptible to the risk of financial crises?

To address this question, I develop a model in which the key mechanism is a build-up of financial fragility in the banking sector during periods of booming productivity growth, when expected returns on banks' long-term investments are high. Building on the seminal work of Diamond and Rajan (2000, 2001b), part of banks' external financing in the model comes in the form of short-term demandable deposits. Deposits

[1] See Reinhart and Reinhart (2010); Mendoza and Terrones (2008, 2014); Schularick and Taylor (2012); Reinhart and Reinhart (2010); Krishnamurthy and Muir (2020).

[2] See Cerra and Saxena (2008); Reinhart and Reinhart (2010); Reinhart and Rogoff (2009, 2014); Krishnamurthy and Muir (2020); Queralto (2020).

[3] For example, different strands of literature emphasized either moral hazard due to implicit government guarantees (Farhi and Tirole, 2012; Chari and Kehoe, 2016), overoptimism (Bordalo et al., 2018), or pecuniary externalities operating through borrowing constraints (see e.g. Lorenzoni (2008)).

[4] See, for example, Queralto (2020); Anzoategui et al. (2019); Guerron-Quintana and Jinnai (2014); Garcia-Macia (2017).

discipline bankers against trying to misappropriate their clients' funds by exposing misbehaving banks to runs, and thus allow them to increase financing. In normal times, banks effectively mix deposits and equity financing to maximize discipline, while at the same time avoiding the risk of destructive runs. However, when the long-run productivity prospects become sufficiently strong, bankers' moral hazard problem becomes worse. This prompts them to rely heavily on deposits, since their ability to raise financing through other means deteriorates. Consequently, it is deposit financing that enables a boom in credit, asset prices, and investment. If future aggregate growth prospects deteriorate prematurely, however, the value of banks' assets falls, rendering them insolvent, and instigating systemic bank runs. The model thus implies that a long period of slow subsequent growth is not just the result of the crisis, but also its cause.

More specifically, banks in the model have access to productive long-term projects that require capital investments. The long-run payoff of banks' projects is uncertain, and can be either high or low, depending on the state of aggregate productivity. Capital is supplied by households, who cannot use it as productively as banks. Banks purchase capital on a competitive market, using own funds and financing they raise from households.

While ideally banks would prefer to finance their uncertain long-term investments with outside equity or similar flexible state-contingent instruments, their ability to issue equity is constrained by moral hazard and limited commitment problems. Specifically, banks can take actions that irreversibly reduce the ability of external investors to collect repayments from them in the future.[5] But because equity investors cannot commit to act against their best collective interest, they cannot credibly threaten to punish and liquidate opportunistic banks that attempt to reduce future repayments. Together, these frictions restrict the amount of equity financing that banks can raise.

Banks can, however, increase their financing by issuing deposits that are demandable at any time and feature a sequential service (aka first-come first-served) property, as in Diamond and Rajan (2001b). Demandability and sequential service create a built-in collective action problem among depositors: when faced with a prospect of lower-than-promised future repayment, each individual depositor finds it optimal to run and demand repayment in full immediately, hoping to be among the first in line to do so. Deposits thus act as a disciplining device, since depositors will run and liquidate bank's assets whenever the bank is unable to guarantee them full repayments in the future, even if it is not in their best collective interest. Anticipating this, the bank has strong incentives to maintain high pledgeability of future returns.

Let us now turn to the central results of the paper. I show that in normal times, when low productivity growth is a fairly likely outcome, banks are careful: they limit

[5]This assumption captures banks being able to create obscure and opaque corporate structures, reduce transparency about nature of their investments, maintain weak accounting standards etc.

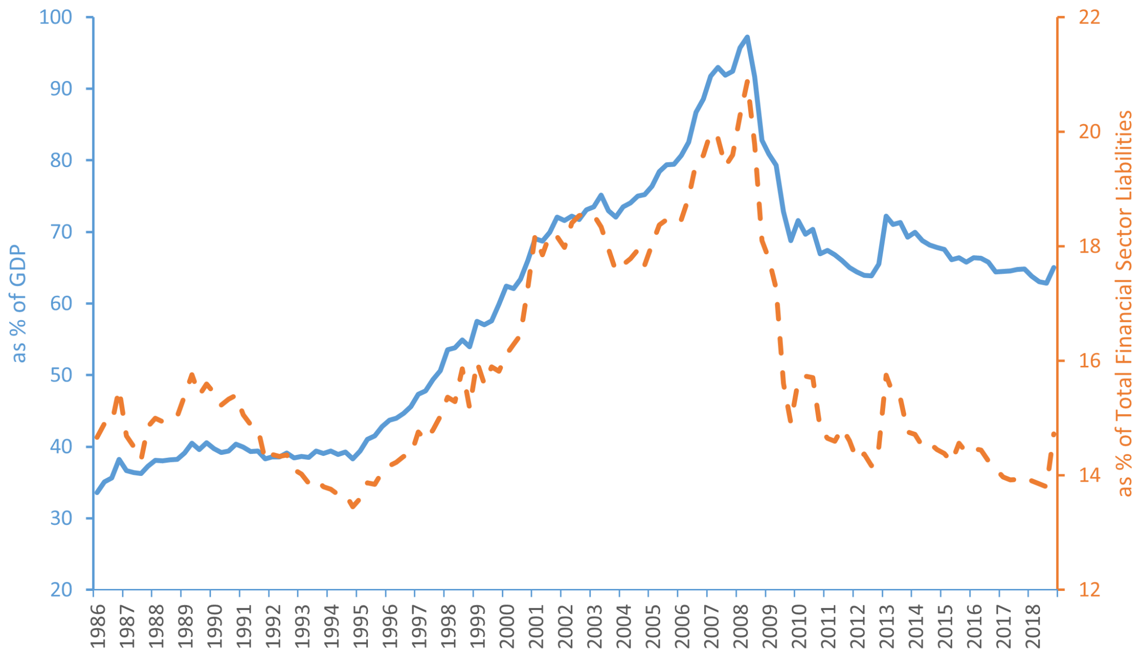


Figure 1: Runnable liabilities in the US before and after the 2008 Crisis

Notes. The data on aggregate runnable liabilities comes from Bao et al. (2015). It includes uninsured deposits, repurchase agreements, securities lending, commercial paper, money market mutual funds shares, variable-rate demand obligations, federal funds borrowed, and funding agreement backed securities. The data on total financial sector liabilities is from the Board of Governors of the Federal Reserve System.

the level of deposits to what they can repay in both high and low productivity states, and issue an incentive-compatible amount of risky outside equity against the high state. When the probability of the high state increases, however, the ability of safe banks to issue outside equity deteriorates, because banks have a greater incentive to squeeze the equity investors' share of the high state returns. Therefore, when the high productivity state becomes sufficiently likely, banks find it optimal to switch to a risky strategy and finance themselves exclusively with deposits, which they would be unable to repay if the low aggregate state materialises. This allows banks to increase the total amount of financing in a boom, and leads to higher equilibrium asset prices and investment. On the dark side, however, high deposit financing introduces financial fragility, and there is a systemic banking crisis in the low productivity state.

Note that the term "deposits" in this paper refers not only to classic bank deposits, but is a metaphor for a wide class of short-term demandable instruments that are capable of creating a collective action problem between investors. For example, Gorton and Metrick (2012) and Martin et al. (2014) argue that it was the run on the repo market that was at the heart of the 2008 Financial Crisis. Similarly, Covitz et al. (2013) document runs on asset-backed commercial paper (ABCP) in 2007. Consistent with the theory developed here, Figure 1 shows that there was a significant accumulation of "runnable" liabilities in the US financial system prior to 2008. The beginning of this build-up coincided with a remarkable resurgence in the productivity growth in the US in the mid-1990s (see Appendix A for details). Then, during the Financial Crisis, there was a large and rapid fall in the outstanding runnable liabilities as many of them, well, were run.

I briefly consider policy implications of the model. The mechanism illustrated in this paper cautions against relying solely on ex-ante preventive policies like leverage restrictions. While a strict macroprudential limit on banks' depositary liabilities can eliminate crises, it also limits investment at times of high productivity growth, precisely when it is the most desirable. Ex-post government bailouts, on the other hand, can save the banking sector from a collapse and improve welfare *in a crisis*. However, in line with much of the literature,[6] anticipation of unrestrained bailouts can completely destroy market discipline of deposits, leading to an inefficient level of investment and hurting welfare. Ex-post interventions are thus only desirable if the government can commit to help banks in case of a systemic banking crisis driven by worsening fundamentals, but not to interfere when the market carries out discipline against opportunistic banks.

Of course, even well-targeted ex-post bailouts deal with symptoms rather than the underlying problem: inability of banks to credibly pledge future returns to investors. This motivates the question: should the banks be monitored during booms to mitigate agency problems and enforce repayments to outside investors? Under simple assumptions about costs and benefits of supervision, I find that the regulator will indeed sometimes supervise banks, increasing their financing and eliminating crises, thus improving welfare. However, the regulator may optimally choose to withdraw and let banks rely on discipline provided by deposits when the confidence in high productivity growth is particularly strong. This potentially sheds light on why, for example, regulation and supervision of the financial sector in the US appear particularly lax in the years preceding the 2008 Financial Crisis.

II. Related literature

This paper builds on Diamond and Rajan (2001b) who, in a similar vein to the seminal paper of Calomiris and Kahn (1991), strive to explain banks' reliance on short-term demandable deposits that expose them to runs. In these papers, demandable deposits with the first-come first-served feature align incentives and preclude banks from acting against depositors' interests. Similarly to the present work, Diamond and Rajan (2000) also analyse optimal bank capital structure under uncertainty. The contribution of my paper is incorporation of this mechanism in the context of a simple macroeconomic model with aggregate uncertainty. I then demonstrate how systemic fragility of the

[6]Prominent papers include Farhi and Tirole (2012), Chari and Kehoe (2016), among many others.

banking sector increases in order to facilitate financing of investments during periods of strong economic growth, thus explaining why financial crises tend to follow credit booms.

The underlying idea that productivity underpins credit booms and financial crises can be traced back to at least Fisher (1933). Closely related recent studies include Gorton and Ordoñez (2019), L'Huillier et al. (2022), Boissay et al. (2016). While in these models crises are certain and perfectly predictable during booms, in my model crises are low probability events, consistent with evidence on agents' expectations.

My paper also contributes to the literature on bank runs. The literature can be loosely divided into "illiquidity" and "insolvency" camps, although the line between the two issues is thin in practice (Thakor, 2018). According to the first camp, rooted in the seminal work of Diamond and Dybvig (1983), runs are panic-based phenomena, and are essentially coordination failures in a world with multiple equilibria. The present paper belongs to the second camp, which seeks to explain bank runs as driven by the insolvency risk of banks due to weak economic fundamentals, consistent with a number of empirical findings.[7] One prominent example of this literature is Allen and Gale (1998), who build a model in which partial bank runs are an efficient way to improve aggregate risk sharing between early and late consumers. In a series of recent notable papers by Gertler and Kiyotaki (2015); Gertler et al. (2019); Martin et al. (2014), while bank runs are still mostly panic-based, run equilibria are only possible when the fundamentals are sufficiently weak. In contrast to the present paper, however, the above papers do not account for why banking crises tend to follow credit booms, or why banks finance themselves with deposits or repos.

In a notable recent paper, Gertler et al. (2020) develop a model in which banking panics are preceded by credit booms, but not all credit booms are "bad", i.e. result in crises. While this is similar to my mechanism, I differ from them in that runnable deposits constitute part of banks' optimal capital structure in my model, and deposit financing endogenously rises in booms. The theoretical advantage of the present framework is that banks cannot eliminate the problem by resorting to a temporary ban on withdrawals in case of a run. While suspension of convertibility would give banks time to renegotiate their liabilities, it would destroy any ex-ante discipline created by deposits.

[7] For example, Calomiris and Gorton (1991) document that banking crises are not random events, but are accounted for by bad macroeconomic news and high insolvency risk of banks; Calomiris and Mason (2003) show that most bank failures during the Great Depression can be well explained by fundamentals, leaving only limited role to panics; Thakor (2018) surveys empirical evidence and concludes that the 2008 Financial Crisis was likely an insolvency risk crisis, rather than an illiquidity crisis.

III. The model

III.A. The framework

There are three periods: $t \in \{0, 1, 2\}$. I will also refer to period 2 as the "long run". There are two types of agents in the economy: households and bankers. There are two types of goods: consumption good, which is the numeraire, and physical capital. Capital is used to produce consumption good, as described below, but cannot be consumed itself. Lastly, all agents are risk-neutral, and, for clarity, I assume that there is no discounting and everyone consumes in period 2.

In period 0, banks can invest in projects that mature and pay off in the long run, i.e. period 2. Banks are scalable, and perfectly diversify away any idiosyncratic projects' risks. Each project requires an investment of one unit of physical capital, which bankers need to purchase on a competitive market. For simplicity, I assume that all capital is owned by households in period 0, and the aggregate supply of capital is fixed and normalised to 1. In Appendix B, I relax the assumption of fixed capital stock and allow for variable aggregate investment.

Each bank starts with a positive endowment of consumption good, which we will call net worth.[8] In what follows, I assume that aggregate net worth of the banking sector N is sufficiently scarce such that banks need to raise additional financing from households to finance their capital purchases. Households, on the other hand, are not constrained in their ability to provide financing to banks in period 0 (or, as explained later, purchase the banking sector's capital in case of liquidation in period 1).[9] The financial market is competitive, and the households break-even in expectation. Figure 2 illustrates the flow of resources in period 0.

There is undiversifiable aggregate uncertainty about the long-run productivity that pertains to all banks' projects. Specifically, there are two possible aggregate productivity states $s \in \{H, L\}$, with respective ex-ante probabilities π and $1 - \pi$. In the high productivity state H, the average return per project is R^H in period 2, whereas in the low productivity state L it is R^L, with $R^L < R^H$. The state is realised in period 1 and is observed by all agents. However, it takes time until period 2 to formally verify the aggregate state. Therefore, financial contracts cannot be state-contingent in period 1.

Projects are bank-specific, and other banks or households cannot run them after the investment is made in period 0.[10] This reflects the idea that banks are relationship

[8] The results would be similar if, instead, we assumed that banks start with a stock of physical capital and liabilities to households.

[9] This, of course, will be always true if the households have deep pockets. This would also be the case, however, even if households start with zero endowments if we assume that all transactions happen instantaneously. In that case, household sector effectively ends up lending capital to banks in exchange of promises of future payments (or, conversely, accepting physical capital in lieu of repayments when banks liquidate projects).

[10] This assumption of extreme project illiquidity ensures that no banks bet on buying projects cheaply

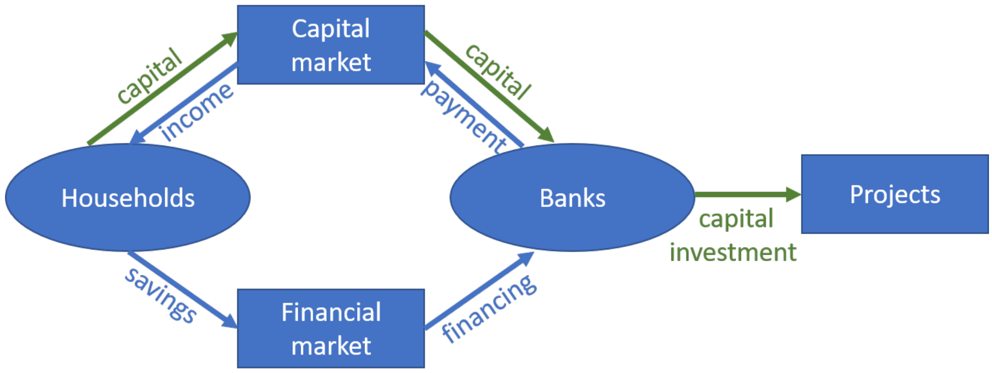


FIGURE 2: Flow of resources in period 0

lenders who possess superior information about their borrowers.[11] Banks can, however, liquidate their projects for Λ per project in period 1, before they mature. Specifically a liquidated project's capital is extracted in full and sold on a competitive secondary capital market back to the households. Households, in turn, employ this capital in a traditional sector that yields Λ per unit of investment. Liquidation is always less profitable than leaving capital with the bank and letting the project mature: $\Lambda < R^L$.[12]

III.B. The frictionless benchmark

It is instructive to describe the first-best outcome in this model. Absent agency problems (introduced below), it is never desirable to liquidate banks' projects in period 1. Any financing that banks raise from households, therefore, takes the form of long-term, state-contingent liabilities that are repaid in period 2 when projects mature.

Consider a bank with net worth n. If the bank raises additional funds b, it can start $(n+b)/Q$ projects, where Q is the market price of capital. Since the bank's creditors break-even in expectation, the expected value of future repayments is simply b, and the bank's expected return is:

$$\frac{n+b}{Q}\mathbb{E}[R^s]-b,$$

where $\mathbb{E}[R^s]=\pi R^H+(1-\pi)R^L$ is the expected return on banks' projects. Clearly, as long as $Q \neq \mathbb{E}[R^s]$, banks are either willing to expand borrowing to infinity, or are not willing to invest into projects at all, neither of which can be the equilibrium. The first-best equilibrium thus requires $Q^* = \mathbb{E}[R^s]$, so that the households' capital sells at fair value. Intermediation is costless and efficient, and there is no capital invested in the traditional

off distressed counterparties in a crisis, and considerably simplifies the analysis.

[11] Although in this model I assume for simplicity that banks manage their projects directly, the setup can also be interpreted as one in which there are entrepreneurs who run the projects, but (a) entrepreneurs are penniless and need to raise external financing; (b) households are unable to finance the entrepreneurs directly due to (unmodelled) agency problems, whereas banks have the expertise to resolve them; (c) entrepreneurs compete for bank funding, so all project returns go to banks.

[12] The assumption that the liquidation value does not depend on the state is not critical for the results.

sector.

IV. Agency problems and financial contracts: equity vs. deposits

In this section I outline the agency problems that banks suffer from, and explore the instruments – equity and deposits – that they can use to raise financing from households. As we shall see, banks financed by equity are never inefficiently liquidated. However, the ability of banks to finance their capital investments using only outside equity is limited. This is because, after investments are made, bankers can take actions that reduce the ability of outsiders to collect repayments from them in the future, while equity investors cannot commit to punish such misbehaving banks. As in Diamond and Rajan (2001b), banks can overcome this problem by issuing demand deposits, which effectively commit depositors to run and liquidate banks that cannot repay them in full. While deposits bring about the needed discipline, they can also lead to socially undesirable bank failures.

IV.A. Financial frictions

The first-best allocation described in the previous section may be unattainable in equilibrium due to the following financial frictions. Banks suffer from limited pledgeability of future returns: at most fraction $\theta \in (0,1)$ of banks' returns is collectable by external claimholders in period 2. What is worse, after the investments in period 0 are made, banks can take actions that further irreversibly reduce the ability of investors to collect from them in the future. Formally, a bank can set the "recovery" rate $\tilde{\theta} \in [0,\theta]$, such that at most fraction $\tilde{\theta}$ of its projects' output can be collected by outsiders in period 2. This assumption reflects the ability of banks to create obscure and opaque corporate structures, limit transparency about the nature of their investments, stick with weak accounting standards etc. A bank's choice of $\tilde{\theta}$ is observable by its investors in period 1, but not verifiable and cannot be included in a contract. Bankers cannot commit to maintain high $\tilde{\theta}$ if it is not in their best interest. In addition, there is also limited commitment on behalf of households, precluding banks from being able to purchase insurance against the aggregate state.

Financial contracts can give investors control rights to liquidate banks' assets. I assume that it is always possible to force banks to liquidate projects and pay proceeds to the investors in period 1, irrespective of their chosen $\tilde{\theta}$'s. Without the liquidation rights, investors would not be able to get any repayments at all, since banks would always set $\tilde{\theta} = 0$; banks thus would not obtain any financing from households in the first place. To keep the model interesting, I assume $\theta R^L > \Lambda$, so that even in the low state investors may receive more than the liquidation value of assets if the bank maintains high $\tilde{\theta}$.

IV.B. Equity financing: well-coordinated investors

One financing option that banks have is to issue long-term state-contingent liabilities, which we will call (outside) equity for convenience. The general equity contract we consider specifies control rights to liquidate assets in period 1, as well as promised state-contingent payouts in period 2. If investors are not repaid in full in period 2, they are entitled to everything that can be recovered from the bank.[13]

Crucially, equity investors are assumed to be well-coordinated and to always act as one to maximize their collective payoff. As in Diamond and Rajan (2000), they cannot commit ex ante to act against their best collective interest ex post. As the result, equity investors will never liquidate assets of a bank if they stand to gain more from letting it continue running until period 2, even if the bank is not able to repay the promised amount in full. This lenience is a double-edged sword. On the one hand, it ensures that equity-financed banks will never get inefficiently liquidated while they continue to generate value for their investors. However, it also allows banks to squeeze their investors by reducing their ability to collect repayments from them in the future. As a consequence, the ability of banks to finance themselves with equity is limited:

LEMMA 1. *The amount of financing that a bank can raise by only issuing equity is at most a fraction* $\frac{\Lambda}{R^L}$ *of the expected returns on its assets.*

This result follows because investors anticipate that the bank can always set $\tilde{\theta} = \Lambda/R^L < \theta$ after selling equity, without risking liquidation. As investors can still collect Λ per project in period 2 even if the low state materialises, they will never liquidate.[14]

IV.C. Deposits: creating collective action problem among investors

The central problem with equity is that well-coordinated investors cannot make credible threats to punish and liquidate banks that try to squeeze them, since this may not be in their best interest ex post. In the seminal paper, Diamond and Rajan (2001b) argue that the use of "harder" claims like demandable deposit contracts can provide the necessary discipline for the banks. The idea is that deposit contracts which promise repayments on demand (forcing banks to liquidate assets in period 1, if necessary), and feature the sequential service (aka first come, first served) property, can break the coordination between investors and make it impossible for the banks to reduce their liabilities on deposits with impunity. This is because there is a built-in collective action problem among depositors: whenever a bank is unable to guarantee full repayments to

[13]While this contract looks more like state-contingent long-term debt rather than equity, the results are the same if instead we consider a more traditional but restrictive equity contact that entitles investors to a fixed share of bank's returns in period 2.

[14]Note that the bank may have incentives to set $\tilde{\theta}$ even lower, thus risking liquidation in the low state and further limiting financing available ex ante.

all of them, each individual depositor has an incentive to run on the bank hoping to be among the first in line and be made whole, rather than accept a lower repayment. An attempt to squeeze depositors will thus trigger a run in which all depositors try to withdraw, even when it is in their best collective interest to agree to lower payments.[15] To put differently, deposit contracts commit bank creditors to punish and liquidate opportunistic banks ex ante, even if they would not choose to do so ex post. Anticipating this, banks that are funded by deposits to a sufficiently high extent will have stronger incentives to maintain high pledgeability of their future returns, allowing them to borrow more. I assume that deposit contacts are available, and constitute the only mechanism capable of creating a collective action problem among banks' investors in the economy.

Note that because the state cannot be verified until period 2, the face value of deposits in period 1 cannot be made state-contingent. In fact, I assume without loss of generality that banks issue only simple deposits which have their face value fixed across all periods and states; banks then can issue equity against the states in which they are able to repay more than the face value of deposits.[16] Lastly, by virtue of their instant demandability, deposits are always senior to equity: presuming that depositors can run and liquidate assets quicker than equity investors in period 1, equity always gets only the residual value, if any, after all depositors are served.[17]

Crucially, financing with deposits also has a dark side. A bank may be inefficiently run and liquidated even when it does not try to extract rents from its creditors, if the pledgeable value of its assets falls below the value of its depositary liabilities due to unfavourable economic conditions. In this case, it would be desirable for depositors to agree to lower repayments, but deposit contracts are explicitly designed to prevent that from happening. At the same time, I assume that banks cannot fall prey to classic panic bank runs as in Diamond and Dybvig (1983), and solvent banks are always able to obtain financing in period 1.

Figure 3 summarizes the timing of events. Let us now turn to how banks combine equity and deposit financing in equilibrium.

[15] For the detailed formal development of this argument, see section III.D, and also the proof of Lemma 2 in the Appendix of Diamond and Rajan (2001b).

[16] Firstly, it would be suboptimal for a bank to issue deposits that promise lower payments in period 2 than in period 1 in some states, since depositors would then always run in period 1 in those states. Secondly, while in principle it is possible to make period-2 payments on deposits state-contingent, as long as they exceed the face value of deposits in period 1, this would not make a difference to the total amount of financing that a bank is able to raise. Since a prospect of a reduction in future repayments would not trigger a run as long as depositors get at least as much as the current face value of their deposits, they cannot commit to not renegotiate the portion of period-2 repayments that is in excess of the face value of deposits in period 1. But then depositors are not different from equity investors with respect to those excess repayments.

[17] This is because in period 1 equity investors will always agree to a necessary reduction in their period 2 claims to forestall a run by depositors.

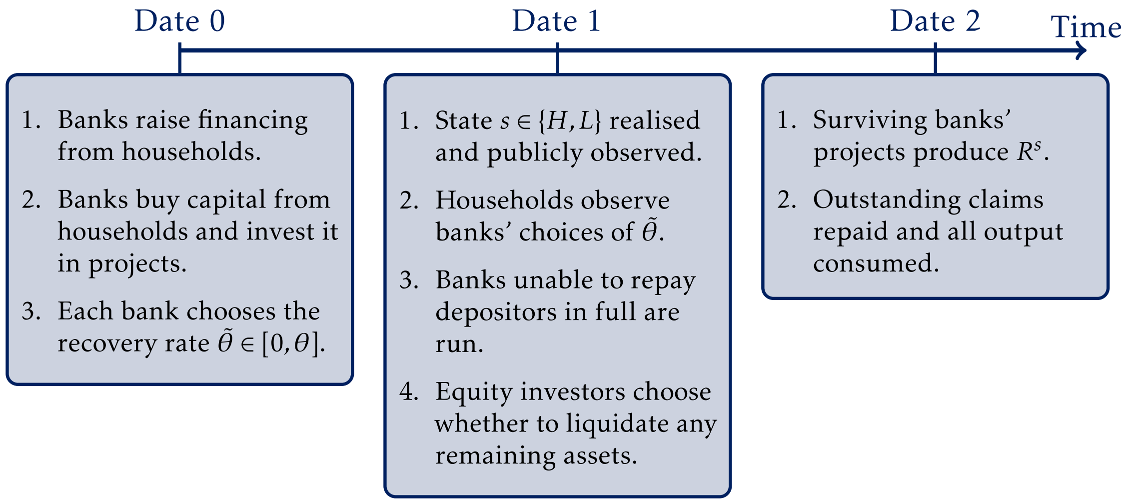


Figure 3: Timeline

V. Banks' choice of financing, equilibrium, and financial crises

In this section we first consider the problem of an individual profit-maximizing bank that takes the market price of capital as given. I show that when the probability of the high productivity state is not too high, the bank always chooses to be safe. It maximizes its payoff by relying on a safe mix of deposits that it can repay in every state of the world and outside equity. However, when the high productivity state becomes more likely, the ability of a bank with safe level of deposits to issue outside equity declines due to moral hazard. When the probability of the high state is sufficiently high, the bank finds it optimal to leverage itself up with deposits beyond what it can repay in the low state, i.e. become risky. This improves discipline and allows the bank to pledge more high-state returns to external investors, but comes at the cost of making a run inevitable in the low state. Second, I show that the model has a unique equilibrium. Similarly to the case of an individual bank, when the probability of the high state is sufficiently high, and the aggregate net worth of the banking sector is sufficiently scarce, all banks choose the risky strategy in equilibrium. This facilitates a boom in asset prices and aggregate investment, but results in a systemic financial crisis if the low state materializes.

V.A. *Banker's problem*

Let us suppose throughout this section that $Q < \mathbb{E}[R^s] \equiv \pi R^H + (1-\pi)R^L$, so banks are financially constrained, earn positive expected rents, and want to expand their operations. Later we will see that this will indeed be the case in equilibrium when the aggregate net worth of the banking sector is sufficiently scarce and the probability of the high state π is sufficiently high. The less interesting case $Q = \mathbb{E}[R^s]$ arises only when the banks' financing constraints do not bind in equilibrium and the first-best outcome is achieved.

Let b be the total amount of financing that a bank with net worth n raises, and so $m \equiv (n+b)/Q$ be the number of projects that the bank starts. Let d be the face value of the bank's deposits, and e^s be the value of the bank's equity in state s in period 2. Without loss of generality, we can assume that e^s is an incentive-compatible amount that the bank actually pays out to equity investors in state s. In period 1 and state s, households would liquidate the bank if $m\tilde{\theta}R^s < \max\{d, m\Lambda\}$, i.e. if either the pledgeable value of the bank's assets is not enough to repay its depositors in full, or is below the liquidation value of the assets.[18]

Depending on its capital structure, the bank can be either *safe* or *risky*. A bank is safe if it is solvent and avoids liquidation of its assets in both high and low aggregate productivity states. A risky bank, on the other hand, gets liquidated by its investors if the low state materialises. Intuitively, the bank's riskiness depends on how much it relies on deposits as the disciplining device. The bank will choose to be safe or risky depending on which strategy yields a higher expected payoff. We can thus look at these two strategies separately, and compare the resulting payoffs in order to determine which one is chosen by the bank.

V.A.1 Safe strategy

Taking as given that it is optimal for the bank to be safe, the bank's maximization problem can be written as:

$$\max \quad v = \pi\left[mR^H - d - e^H\right] + (1-\pi)\left[mR^L - d - e^L\right] \tag{1}$$

$$\text{s.t.} \quad m = \frac{n+b}{Q} \tag{2}$$

$$b = d + \pi e^H + (1-\pi)e^L \tag{3}$$

$$d + e^s \leq m\theta R^s, \; s \in \{L, H\} \tag{4}$$

$$d + e^L \leq \max\{d, m\Lambda\} \tag{5}$$

$$d + e^H \leq \frac{R^H}{R^L}\max\{d, m\Lambda\} \tag{6}$$

$$v \geq \pi\left(mR^H - \max\{d, m\Lambda\}\right) \tag{7}$$

Constraint (3) is the break-even condition for all bank's investors; (4) restricts liabilities of the safe bank to what it can feasibly pay out in each state (if it maintains the recovery rate $\tilde{\theta}$ at the maximum level θ); (5) and (6) place limits on the maximum repayments in the low and high states respectively that the bank cannot reduce simply by lowering $\tilde{\theta}$ to $\max\{d, m\Lambda\}/(mR^L)$, and still avoiding liquidation in both states; lastly (7) is the incentive compatibility constraint which ensures that the safe bank does not benefit

[18]Note that it always optimal for a financially constrained bank to give investors all control rights in period 1 in this model.

in expectation from squeezing equity investors' share of returns in the high state by setting $\tilde{\theta} = \max\{d, m\Lambda\}/(mR^H)$, and accepting the risk of failure in the low state.

Since $Q < \pi R^H + (1-\pi)R^L$, the bank will strive to raise as much external financing as possible, i.e. v is maximized when b is maximized. A quick analysis of the constraints reveals that this is achieved when $e^L = 0$ and $d = m\theta R^L$, i.e. the bank operates with the highest possible safe level of deposits, and issues equity only against the high state. This is intuitive, because deposits provide discipline, and therefore relax constraints (5)-(7). The solution to the safe bank's profit maximization problem then depends on which one of the constraints (6) and (7) binds, which in turn depends on the probability of the high state π. Formally, it is straightforward to show that:

Lemma 2. *A safe bank always operates with $e^L = 0$ and $d = m\theta R^L$, i.e. the largest safe level of deposit liabilities, in order to maximise discipline. Let*

$$\dot{\pi} \equiv \frac{(1-\theta)R^L}{\theta(R^H - R^L) + (1-\theta)R^L} \in (0,1). \tag{8}$$

- *When $\pi \leq \dot{\pi}$, constraint (6) binds, constraint (7) is slack, and the bank is able to pledge the maximum fraction θ of expected future returns to external investors. It raises $\theta\mathbb{E}[R^s]$ of external funds per project, starts $m_s = n/(Q - \theta\mathbb{E}[R^s])$ projects, and its expected payoff is $v_s = m_s(1-\theta)(\pi R^H + (1-\pi)R^L)$.*
- *When $\pi > \dot{\pi}$, constraint (7) binds, and the bank can only raise $\pi\theta R^L + (1-\pi)R^L$ of external financing per project, while remaining safe. The bank runs*

$$m_s = \frac{n}{Q - \pi\theta R^L - (1-\pi)R^L} \tag{9}$$

projects, and earns the expected payoff

$$v_s = m_s\pi\left(R^H - \theta R^L\right). \tag{10}$$

Corollary 1. *The ability of a safe bank to finance its projects externally falls as the high state becomes more likely.*

The intuition for this result is as follows. As the high state becomes more likely, the banker's moral hazard becomes worse. He has a stronger incentive to squeeze equity investors in order to get a higher share of the projects' returns in the high state, so the amount of outside equity the bank can issue falls. The ability of a safe bank to use deposits to discipline itself is limited by the fact that it must be able to repay depositors in full in all states. The solid line in Figure 4 illustrates the fraction of expected projects' returns pledgeable by a safe bank to households as a function of the probability of the high state π.

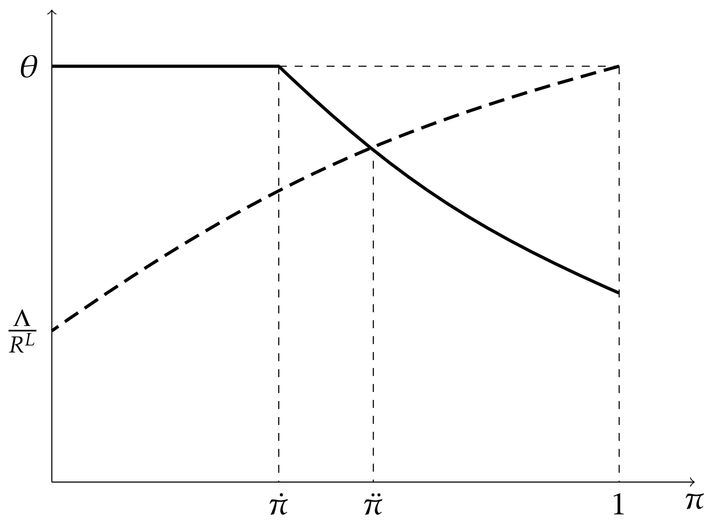


Figure 4: Fraction of expected projects' value that safe (solid line) and risky (dashed line) banks can pledge to external investors

V.A.2 Risky strategy

This brings us to the second, risky, strategy that the bank may pursue when the high state is sufficiently likely. Namely, the bank can accept that it will fail in the low state, but use deposits to raise more financing against the high state, and therefore invest in more projects. Taking as given that the bank has $d > m\theta R^L$ and fails in the low state, its problem becomes:

$$\max \quad v = \pi\left(mR^H - d\right) \tag{11}$$

$$\text{s.t.} \quad m = \frac{n+b}{Q} \tag{12}$$

$$b = \pi d + (1-\pi)m\Lambda \tag{13}$$

$$d \leq m\theta R^H \tag{14}$$

Constraint (13) is the break-even condition of the bank's depositors, who run and liquidate the bank in the low state;[19] (14) limits deposit liabilities of the risky bank to what it can feasibly repay in the high state. The risky bank issues no outside equity and $e^H = 0$, since otherwise it could squeeze equity investors out by picking the recovery rate at $\tilde{\theta} = d/(mR^H)$, while still avoiding the run and liquidation in the high state.

Combining (11) and (13), it is easy to see that for the risky strategy to earn a positive expected payoff, and thus for the bank to even consider it over the safe one, it must be that $Q < \pi R^H + (1-\pi)\Lambda$, i.e. the capital price should not exceed the value of the projects (accounting for the liquidation in the L state). When it is the case, the bank's payoff

[19] Note that this is the aggregate break-even condition. In case of a run, some individual depositors get the full face value of their deposits, and some get nothing; but the total payout equals to the liquidation value of the bank's assets, $m\Lambda$.

is again maximized when external financing b is maximized. This is achieved when the bank issues the maximum level of deposit liabilities, $d = m\theta R^H$. This way the bank commits to repay the maximum pledgeable fraction θ of the returns in the high state.

Lemma 3. *The solution to the risky bank's problem is achieved when constraint* (14) *binds and the bank promises the maximum pledgeable fraction θ of returns in the high state to depositors. This allows the bank to raise external financing $\pi\theta R^H + (1-\pi)\Lambda$ per project, start*

$$m_r = \frac{n}{Q - \pi\theta R^H - (1-\pi)\Lambda} \tag{15}$$

projects, and earn the expected payoff

$$v_r = m_r \pi (1-\theta) R^H. \tag{16}$$

The risky strategy allows the bank to raise more external financing against the high state. Therefore, as the high state becomes more likely, the ability of a risky bank to finance itself externally improves. The dashed line in Figure 4 shows the amount of external financing that a risky bank can raise, normalised as a fraction of the first-best expected projects' value $\mathbb{E}[R^s]$ (to enable the comparison with the safe case). As π approaches 1, this fraction approaches the maximum feasible value θ.

V.A.3 Safe or risky?

So when will the bank adopt a high, risky level of deposits?

Lemma 4. *The necessary conditions for a bank to even consider the risky strategy are $Q < \pi R^H + (1-\pi)\Lambda$ and $\pi > \ddot{\pi}$, where*

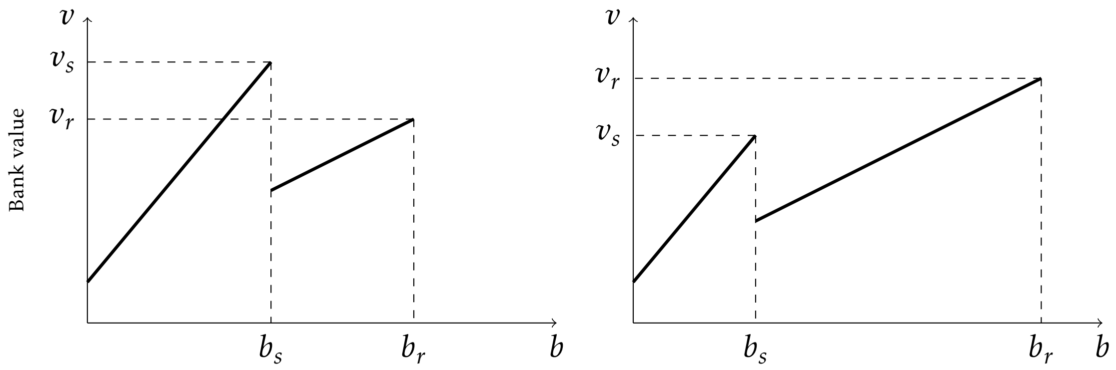
$$\ddot{\pi} \equiv \frac{R^L - \Lambda}{\theta(R^H - R^L) + R^L - \Lambda} \in (\dot{\pi}, 1). \tag{17}$$

Then, the bank strictly prefers the risky strategy to the safe strategy iff $Q < \overline{Q}$, where:

$$\overline{Q}(\pi) \equiv \pi R^H + (1-\pi)\Lambda - \frac{(1-\pi)(1-\theta)(R^L - \Lambda)R^H}{\theta(R^H - R^L)}, \tag{18}$$

i.e., ceteris paribus, either (a) π is sufficiently high given Q; or (b) Q is sufficiently low given π.

When the probability of the high state π is below the $\ddot{\pi}$ threshold, the bank can obtain more external financing by following the safe strategy, as can be seen in Figure 4. Hence there would no benefit at all of leveraging up with the risky level of deposits.



(a) Bank chooses safe strategy

(b) Bank chooses risky strategy

FIGURE 5: Bank's value as a function of external financing

When $\pi > \ddot{\pi}$, the bank trades off the ability to raise more financing and run more projects enabled by the risky strategy vs. losses due the bank run in the low state. Figure 5 depicts the banker's expected value as a function of the bank's external financing b, taking the market price of capital Q as given. The bank's value increases linearly up until b_s (the maximum amount of financing that the bank can raise while remaining safe). At that point the value drops discontinuously: to raise more external financing than b_s, the bank must lever itself up with deposits beyond what it can repay in the low state, implying a run and inefficient liquidation of its projects in the L state. The value then continues to increase linearly in b up until b_r (the maximum amount that the risky bank can borrow), although now with a shallower slope, reflecting the lower value of each project. The bank's problem reduces to simply choosing between the two peaks: v_s (playing it conservatively) and v_r (betting on the high productivity, but suffering a run if it does not materialise), whichever yields a higher value. Comparing these values given in (10) and (16), the second claim in the lemma follows.[20] The risky strategy is more appealing when the high productivity state becomes more likely and/or the capital becomes cheaper. The latter is because the additional financing that the bank is able to obtain translates into more additional projects when Q is low.

V.B. *Equilibrium and financial crises*

Equilibrium requires that financial and capital markets clear in period 0. The key equilibrating variable is the price of capital Q: the demand for capital from banks that take Q as given must equal the supply of capital by households. At the same time, the banks' demand for capital depends on the amount of external financing they obtain, which, as we have seen, also depends on Q.

[20] In the lemma, I implicitly use the fact that in equilibrium we will necessarily have $Q > \max\left\{\pi\theta R^H + (1-\pi)\Lambda, \pi\theta R^L + (1-\pi)R^L\right\}$, otherwise banks could make infinite profits and would demand infinite amount of capital.

We say that there is a risk of a financial crisis if all or a substantial fraction of banks choose the risky strategy in period 0, and are run and liquidated in the L state. The central result of this paper, formally established in the proposition below, is that such risk indeed emerges in equilibrium during booms, when the probability of the high productivity state is large.

Proposition 1. *There is an essentially unique equilibrium in this economy (with respect to aggregate outcomes),*[21] *which depends on the probability of the high state π. Suppose $N < (1-\theta)R^H$, i.e. the aggregate net worth of the banking sector is sufficiently scarce. Then there are thresholds*

$$\underline{\pi} = \frac{N\frac{\theta(R^H-R^L)}{R^H-\theta R^L} + R^L - \Lambda}{\theta(R^H - R^L) + R^L - \Lambda} \quad and \quad \overline{\pi} = \frac{N\frac{\theta(R^H-R^L)}{(1-\theta)R^H} + R^L - \Lambda}{\theta(R^H - R^L) + R^L - \Lambda},$$

with $\ddot{\pi} < \underline{\pi} < \overline{\pi} < 1$, such that:

a. *If $\pi \geq \overline{\pi}$, there is a unique equilibrium in which all banks are financed with the high, risky level of deposits. This facilitates a credit and asset price boom in period 0: banks borrow $\pi\theta R^H + (1-\pi)\Lambda$ per project, and drive the equilibrium price of capital up to*

$$Q^* = N + \pi\theta R^H + (1-\pi)\Lambda. \tag{19}$$

There is a systemic financial crisis in which all banks fail if the low state materialises.

b. *If $\pi \in (\underline{\pi}, \overline{\pi})$, a fraction $\zeta^* \in (0,1)$ of banks (measured by the net worth)*[22] *end up choosing the risky level of deposits in equilibrium, while the remaining banks are safe. The equilibrium price of capital is $\overline{Q}$ given in (18). The fraction ζ^* of banks that fail in the low state is given by*

$$\zeta^* = \frac{(1-\theta)R^H}{N\theta(R^H - R^L)}[Q^* - \pi\theta R^L - (1-\pi)R^L - N],$$

which is an increasing function in the ex-ante probability of the high state: $\zeta_\pi^{*\prime} > 0$.

c. *If $\pi \leq \underline{\pi}$, all banks are safe, and never fail in equilibrium.*

Proof. See Appendix C. *Q.E.D.*

---

[21] While there may be multiple equilibria (with respect to strategies of individual banks), notably in cases (b) and (c), the essential uniqueness here means that these equilibria are always equivalent in terms of the aggregate outcomes, e.g. the amount of credit, the price of capital, and the fraction of banks that end up being risky.

[22] Here we require a technical assumption that banks are atomistic, such that such split is always possible for any $\zeta \in (0,1)$.

Note that the equilibrium capital price and aggregate bank financing in cases (a) and (b) are higher than what they would be if all banks chose to be safe.[23] Households thus receive higher payment for their capital, while banks extract lower rents for their intermediation services. While so far we have assumed that the aggregate capital stock is fixed, it is likely that booms in the asset price and credit occur simultaneously with booms in the aggregate investment. Indeed, I show this in Appendix B, where I formally extend the above results to the case in which the aggregate capital investment can vary.

The condition $N < (1-\theta)R^H$ is realistic. Assuming otherwise would imply that the banking sector has enough own funds to bid up the capital price to the full value of the investment projects even when the high state is certain (and thus banks can never earn positive rents). But it is unlikely that banks would have accumulated that much net worth during normal times when the probability of the high state is lower.

To summarise the results, during booms in productivity, the worsening moral hazard problem induces banks to rely exclusively on deposit financing, which facilitates a boom in asset prices, credit, and aggregate investment. If the anticipated growth fails to materialise, however, the excessive deposit financing leads to a systemic financial crisis.

VI. Implications

In this section I discuss the implications of the theory presented above. The ex-ante leverage restrictions and ex-post bailouts may hurt welfare by stifling investment and destroying discipline, respectively. However, the regulatory supervision of banks that enhances their transparency and governance can sometimes eliminate crises and improve welfare.

VI.A. Ex-ante leverage restrictions

While financial crises in the model are destructive and undesirable ex post, equilibria that feature crises are constrained efficient ex-ante. There are no externalities that agents do not take into account in this simple model. Even though a growing reliance of banks on deposit-like short-term financing introduces the risk of a systemic banking crisis, it also allows to increase capital investment when it is the most valuable: during booms with strong prospects of high future returns.

It is easy to see that sufficiently hard restrictions on ability of banks to leverage with deposits can always preclude financial crises in the present setting. However, this will not be Pareto improving: restricted leverage also means lower capital prices, and lower income for households, who are capital suppliers. Leverage restrictions can lead to

[23]Specifically, safe banks would be able to raise only $\pi\theta R^L + (1-\pi)R^L$ per project by Lemma 2, and the price of capital would be $N + \pi\theta R^L + (1-\pi)R^L$. These quantities are lower than those given in the Proposition in cases (a) and (b), since $\pi > \underline{\pi} > \ddot{\pi}$ in these cases.

higher profits of the banking sector, since they effectively limit the competition between banks on the capital market, and allow them to earn higher rents per project. Nevertheless, there is a net loss to the society in the form of missed investment opportunities, once variable capital supply is introduced into the model.

More realistic settings would surely incorporate additional social costs of bank failures, which may not be fully internalised by the financial market participants (e.g. fire sales of assets, unemployment, aggregate demand externality etc). The use of leverage restrictions to preclude financial crises could therefore be desirable once these externalities are taken into account. The contribution of the present work is simply to highlight that such policies can also have significant economic costs.

VI.B. Cleaning-up ex post: government bailouts

Unlike the models in which financial crises stem from panics, notably Diamond and Dybvig (1983), in the present model banking crisis is not a result of agents' failure to coordinate on a "good" equilibrium. If growth opportunities deteriorate during a boom, banks are run by depositors because they become insolvent: their liabilities on deposits exceed what they can possibly pledge to repay.

Nevertheless, saving banks in a crisis could in principle be desirable *ex post*, since liquidation of banks' assets destroys wealth for both households and bankers. In practice, government rescue could take many forms: costly government guarantees of problematic banks' debt; lending to troubled banks at below-market interest rates; or simply injecting equity into banks. Arguably, most of these tools were used in the US during the 2008 Financial Crisis.[24]

However, unrestrained government support to the banking sector could cause problems *ex ante*. As in Diamond and Rajan (2001b), the theory suggests that fragile capital structure may be a market solution to the bank discipline problem. Therefore, similarly to the large existing literature on moral hazard induced by implicit government guarantees,[25] expectations of government rescue could completely destroy ex-ante incentives. If creditors expect that the government will use taxpayers' money to bail out troubled banks, they will have no reasons to worry about bankers' conduct and incentives, leading to an overheated credit market, inflated asset prices, and high rent extraction by bankers. Incentives of banks to maintain high pledgeability of returns would be lost, and aggregate investment could be severely distorted.

[24]Despite the fact that Lehman Brothers was allowed to fail because the Fed deemed it insolvent at the time, one can argue that other insolvent banks were effectively bailed out. Notwithstanding claims that the US government has made money on assisting troubled banks, this does not appear to be true once one takes into account the fair economic value of government's assistance. Lucas (2019) estimates the direct cost of bailouts to the US government to be in excess of £500bn, or 3.5% of GDP.

[25]Some of the most prominent papers on the topic, among others, are Chari and Kehoe (2016); Farhi and Tirole (2012).

Government bailouts would thus only be desirable ex ante if they are carried out in a way that preserves the market discipline. This requires that the government is able to commit to bail out banks only in crises driven by significant aggregate shocks outside of banks' control, but let fail any banks that try to extract rents from their creditors or taxpayers. How to generate such a commitment remains an important question actively discussed in academic literature and policy circles. In essence, this is an institutional design problem of giving proper incentives to the regulator responsible for bailouts. I can only conjecture that the optimal design would require that the regulator (i) is well informed about the aggregate state of the economy, and (ii) is more reluctant to assist banks than a benevolent social planner would be. The latter would ensure that the regulator does not intervene unless there are significant benefits to the society from doing so. This could be achieved by making bailouts legally difficult and politically costly to the regulator.

VI.C. Addressing financial frictions directly

The fundamental issue in the model is the inability of banks to credibly pledge future returns to external investors. The reliance of banks on rigid, demandable deposits during booms, as well as the occasionally resulting financial crises are thus just the symptoms of the underlying problem.

One implication of the theory is the prediction that banking crises are more likely to erupt in developing countries. Many such countries are characterised by weak financial institutions and corporate governance problems, while, at the same time, having less stable political regimes and more volatile economic trends. Empirically, we indeed observe more frequent crises in the developing economies. The second implication is that the best remedy against financial crises is a developed and transparent financial sector with sufficient intermediation capacity to allow banks to refrain from excessive short-term financing – a point also forcefully stressed in Diamond and Rajan (2001a).

We can also ask a more immediate question: should the government monitor banks during booms in order to mitigate agency problems and enforce repayments to outside investors? The answer depends on the feasibility and costs of monitoring. Formally, suppose that by supervising banks, the regulator can eliminate the moral hazard problem completely and ensure that a fraction θ of banks' returns gets repaid. However, active regulatory interference in banks' business is also likely to be costly. I assume that it reduces future returns by a fraction $1-\chi$, such that banks' projects only generate χR^s in state $s \in L, H$. Let us also assume that the regulator is household-friendly, i.e. its sole objective is to maximize the welfare of household, and places zero weight on bankers' profits. Since households supply capital, and always break-even in expectation on the financial market, households' welfare maximization is equivalent to maximizing the

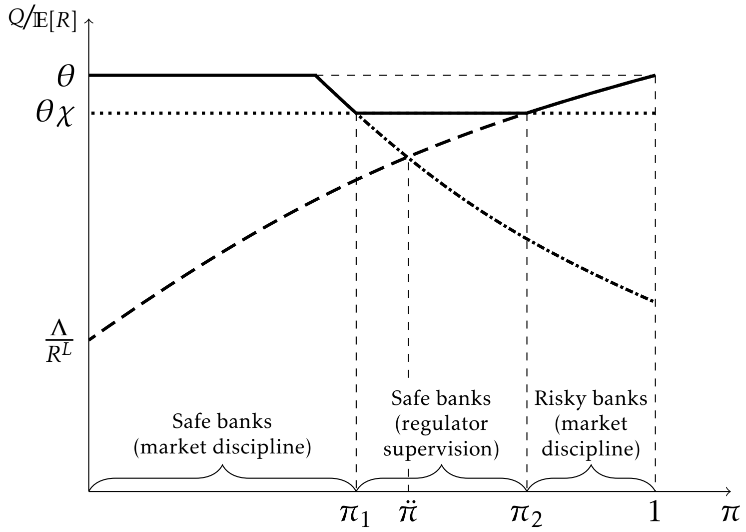


Figure 6: Capital price as a fraction of its first best level: market discipline vs regulatory supervision

capital price. This in turn depends on the amount of financing that banks can sustain. Then we can show that:

Proposition 2. *If the cost of supervision* $1-\chi$ *is sufficiently low, specifically*

$$\chi > \frac{\max\left\{\underline{\pi}R^L + \frac{1-\underline{\pi}}{\theta}R^L, N + \underline{\pi}\theta R^L + (1-\underline{\pi})R^L\right\}}{\underline{\pi}R^H + (1-\underline{\pi})R^L}, \tag{20}$$

there exist π_1, π_2, *with* $\pi_1 < \underline{\pi} < \pi_2 < 1$, *such that banks are supervised iff* $\pi \in (\pi_1, \pi_2)$.

However, there is no supervision otherwise, and the equilibrium features the risk of a financial crisis in the low state when $\pi > \pi_2$.

Proof. See Appendix C. *Q.E.D.*

The main takeaway is that the regulator will indeed actively supervise banks and eliminate the risk of a financial crisis when the probability of the high state becomes moderately high. The regulator, however, will choose to leave banks to the market discipline once the high productivity state becomes very likely. Intuitively, when the probability of the low state becomes small, expected benefit of preventing a crisis falls; the regulatory supervision thus becomes more costly than deposits in forcing repayments from banks in the high state. This result potentially sheds some light on why regulation and supervision of the financial sector in the US appeared particularly lax in the years preceding the 2008 Financial Crisis: On the back of strong confidence in the high productivity growth, the regulators might had been afraid to do more harm than good by interfering.

We can graphically illustrate the proposition in the special case when $N \to 0$ (i.e. when the banking sector has negligibly low aggregate net worth), which is particularly easy to visualise. When banks have very low net worth, the ability to raise external financing becomes their crucial concern in equilibrium, and banks are willing to become risky the moment it allows them to better finance themselves, i.e. when $\pi > \ddot{\pi}$. Figure 6 plots the ratio of capital price Q to its first-best level, $\mathbb{E}[R] = \pi R^H + (1-\pi)R^L$, as a function of the probability of the high state π. The figure depicts three cases: (i) the regulator actively supervises banks (dotted line), (ii) all banks rely on the market discipline and choose the safe strategy (dash-dotted line), or (iii) all banks rely on the market discipline and choose the risky strategy (dashed line). The latter two lines are effectively the same as in Figure 4, because the equilibrium capital price equals to the financing that banks can raise per project when $N \to 0$. In equilibrium under optimal supervision, the price of capital is the highest of these three lines, highlighted by the solid line on the plot. Supervision can achieve greater repayments from banks than the market discipline when $\pi \in [\pi_1, \pi_2]$, and it eliminates financial crises when $\pi \in (\ddot{\pi}, \pi_2]$. However, the regulator prefers to let banks finance themselves with risky deposits when $\pi > \pi_2$.

VII. Conclusion

This paper presents a model in which financial crises result from build-ups of bank leverage and financial fragility during periods of rational optimism, when banks expect high returns on their long-term investments that ultimately do not materialise. The key mechanism is an increasing reliance of banks on deposit financing, which helps them overcome agency problems and increase funding, but also exposes them to runs. The model is consistent with empirical evidence that credit booms and financial crises appear to be preceded by swings in the long-run productivity trends.

Ex-ante leverage restrictions can eliminate crises, but also stifle investment at the time when it is most profitable. On the other hand, unrestrained ex-post bailouts destroy ex-ante incentives. They are therefore desirable only if the government can commit to interfere only when banks require help because of deteriorating macroeconomic fundamentals, but otherwise allow them to fail. How to engender such a commitment is a fascinating direction for future research.

Costly regulatory supervision of banks that enhances their accountability to investors can preclude crises and sometimes improves welfare. However, it may still be optimal for the regulator to let banks rely solely on the market discipline during times when the confidence in high productivity growth is very strong. This potentially explains the apparently lax regulation of the US banks in the run up to the 2008 Crisis.

References

Allen, F. and Gale, D. (1998). Optimal financial crises. *Journal of Finance*, 53(4):1245–1284.

Anzoategui, D., Comin, D., Gertler, M., and Martinez, J. (2019). Endogenous technology adoption and R&D as sources of business cycle persistence. *American Economic Journal: Macroeconomics*, 11(3):67–110.

Bao, J., David, J., and Han, S. (2015). The runnables. FEDS Note, Board of Governors of the Federal Reserve System, Washington.

Benati, L. (2007). Drift and breaks in labor productivity. *Journal of Economic Dynamics and Control*, 31(8):2847–2877.

Boissay, F., Collard, F., and Smets, F. (2016). Booms and banking crises. *Journal of Political Economy*, 124(2):489–538.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2018). Diagnostic expectations and credit cycles. *Journal of Finance*, 73(1):199–227.

Calomiris, C. W. and Gorton, G. (1991). The origins of banking panics: Models, facts, and bank regulation. In *Financial Markets and Financial Crises*, pages 109–174. University of Chicago Press.

Calomiris, C. W. and Kahn, C. (1991). The role of demandable debt in structuring optimal banking arrangements. *American Economic Review*, 81(3):497–513.

Calomiris, C. W. and Mason, J. R. (2003). Fundamentals, panics, and bank distress during the Depression. *American Economic Review*, 93(5):1615–1647.

Cao, D. and L'Huillier, J.-P. (2018). Technological revolutions and the Three Great Slumps: A medium-run analysis. *Journal of Monetary Economics*, 96:93–108.

Cerra, V. and Saxena, S. C. (2008). Growth dynamics: The myth of economic recovery. *American Economic Review*, 98(1):439–457.

Cette, G., Fernald, J., and Mojon, B. (2016). The pre-Great Recession slowdown in productivity. *European Economic Review*, 88:3–20.

Chari, V. V. and Kehoe, P. J. (2016). Bailouts , time inconsistency, and optimal regulation: A macroeconomic view. *American Economic Review*, 106(9):2458–2493.

Covitz, D., Liang, N., and Suarez, G. A. (2013). The evolution of a financial crisis: Collapse of the asset-backed commercial paper market. *Journal of Finance*, 68(3):815–848.

Cummins, J. G. and Violante, G. L. (2002). Investment-specific technical change in the United States (1947-2000): Measurement and macroeconomic consequences. *Review of Economic Dynamics*, 284:243–284.

Dell'Ariccia, G., Igan, D., Laeven, L., and Tong, H. (2012). Policies for Macrofinancial Stability: How to Deal with Credit Booms. Technical report, IMF Staff Discussion Note.

DeLong, J. B. (2003). Productivity growth in the 2000s. In *NBER Macroeconomics Annual 2002, Volume 17*, pages 113–158. MIT Press.

Diamond, D. W. and Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, 91(3):401–419.

Diamond, D. W. and Rajan, R. G. (2000). A theory of bank capital. *Journal of Finance*, 55(6):2431–2465.

Diamond, D. W. and Rajan, R. G. (2001a). Banks, short term debt and financial crises: Theory, policy implications and applications. In *Carnegie-Rochester Conference Series on Public Policy*, volume 54, pages 37–71.

Diamond, D. W. and Rajan, R. G. (2001b). Liquidity risk, liquidity creation, and financial fragility: A theory of banking. *Journal of Political Economy*, 109(2):287–327.

Farhi, E. and Tirole, J. (2012). Collective moral hazard, maturity mismatch, and systemic bailouts. *American Economic Review*, 102(1):60–93.

Fernald, J. G. (2014). Productivity and potential output before, during, and after the Great Recession. In *NBER Macroeconomics Annual 2014, Volume 29*, pages 1–51. University of Chicago Press.

Fisher, I. (1933). The debt-deflation theory of Great Depressions. *Econometrica*, 1(4):337.

Garcia-Macia, D. (2017). The financing of ideas and the Great Deviation. IMF Working Paper 2017/176.

Gertler, M. and Kiyotaki, N. (2015). Banking, liquidity, and bank runs in an infinite horizon economy. *American Economic Review*, 105(7):2011–2043.

Gertler, M., Kiyotaki, N., and Prestipino, A. (2019). A macroeconomic model with financial panics. *The Review of Economic Studies*, 87(1):240–288.

Gertler, M., Kiyotaki, N., and Prestipino, A. (2020). Credit booms, financial crises, and macroprudential policy. *Review of Economic Dynamics*, 37:S8–S33.

Gordon, R. J. (2003). Exploding productivity growth: Context, causes, and implications. *Brookings Papers on Economic Activity*, 2:207–298.

Gorton, G. and Metrick, A. (2012). Securitized banking and the run on repo. *Journal of Financial Economics*, 104(3):425–451.

Gorton, G. and Ordoñez, G. (2019). Good booms, bad booms. *Journal of the European Economic Association*, 18(2):618–665.

Guerron-Quintana, P. A. and Jinnai, R. (2014). Liquidity, trends, and the Great Recession. Working Paper 14-24, Federal Reserve Bank of Philadelphia.

Jorgenson, D. W., Ho, M. S., and Stiroh, K. J. (2004). Will the U.S. productivity resurgence continue? *Current Issues in Economics and Finance*, 10(13):1–7.

Kahn, J. A. and Rich, R. W. (2007). Tracking the new economy: Using growth theory to detect changes in trend productivity. *Journal of Monetary Economics*, 54(6):1670–1701.

Krishnamurthy, A. and Muir, T. (2020). How credit cycles across a financial crisis. Working Paper 3579, Stanford Graduate School of Business.

L'Huillier, J.-P., Phelan, G., and Wieman, H. (2022). Technology shocks and predictable minsky cycles. Working paper.

Lorenzoni, G. (2008). Inefficient credit booms. *The Review of Economic Studies*, 75(3):809–833.

Lucas, D. (2019). Measuring the cost of bailouts. *Annual Review of Financial Economics*, 11(1):85–108.

Martin, A., Skeie, D., and Thadden, E. L. V. (2014). Repo runs. *Review of Financial Studies*, 27(4):957–989.

Mendoza, E. G. and Terrones, M. E. (2008). An anatomy of credit booms: Evidence from macro aggregates and micro data. Working Paper 14049, National Bureau of Economic Research.

Mendoza, E. G. and Terrones, M. E. (2014). An anatomy of credit booms and their demise. In Fuentes, M. D., Raddatz, C. E., and Reinhart, C. M., editors, *Capital Mobility and Monetary Policy*, pages 165–204. Central Bank of Chile, Santiago, Chile.

Paul, P. (2022). Historical patterns of inequality and productivity around financial crises. *Journal of Money, Credit, and Banking*, forthcoming.

Queralto, A. (2020). A model of slow recoveries from financial crises. *Journal of Monetary Economics*, 114:1–25.

Reinhart, C. M. and Reinhart, V. R. (2010). After the fall. In *Macroeconomic Challenges: The Decade Ahead*, pages 17–60. Federal Reserve Bank of Kansas City Economic Policy Symposium.

Reinhart, C. M. and Rogoff, K. S. (2009). The aftermath of financial crises. *American Economic Review: Papers & Proceedings*, 99(2):466–472.

Reinhart, C. M. and Rogoff, K. S. (2014). Recovery from financial crises: Evidence from 100 episodes. *American Economic Review: Papers & Proceedings*, 104(5):50–55.

Schularick, M. and Taylor, A. M. (2012). Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008. *American Economic Review*, 102(2):1029–1061.

Stiroh, K. and Botsch, M. (2007). Information technology and productivity growth in the 2000s. *German Economic Review*, 8(2):255–280.

Thakor, A. V. (2018). Post-crisis regulatory reform in banking: Address insolvency risk, not illiquidity! *Journal of Financial Stability*, 37:107–111.

APPENDICES

Appendix A. Productivity boom and the 2008 Crisis

The Great Recession is a case in point. It is well documented that there has been a marked acceleration in the US productivity growth in the mid 1990s after more than 25 years of sluggish growth, widely attributed to information and communication technologies.[26] Figure A.1, taken from the up-to-date version of Kahn and Rich (2007) two-regime growth model, demonstrates a shift to the high-growth regime around 1997. The gradual recognition of the New Economy prompted many people, including professional economists, to express confidence in high rate of growth for the decades to come.[27]. Figure A.2 shows that medium professional forecasts of 10-year average productivity growth nearly doubled over the 3 years of 1997-2000, and were hardly affected by the dot-com bust. At the time, these projections were not necessarily unreasonable: Since shifts in the trend are infrequent and difficult to predict, the best forecast following an acceleration of productivity may well be to expect that the economy will experience a prolonged period of high growth. The figure also shows that

[26] See Benati (2007); Kahn and Rich (2007); Fernald (2014). In addition, Fernald (2014) shows that although capital deepening played a role, high growth of labour productivity in the late 1990s - early 2000s was driven primarily by TFP.

[27] Among many others were: DeLong (2003); Cummins and Violante (2002); Jorgenson et al. (2004); Stiroh and Botsch (2007); Gordon (2003).

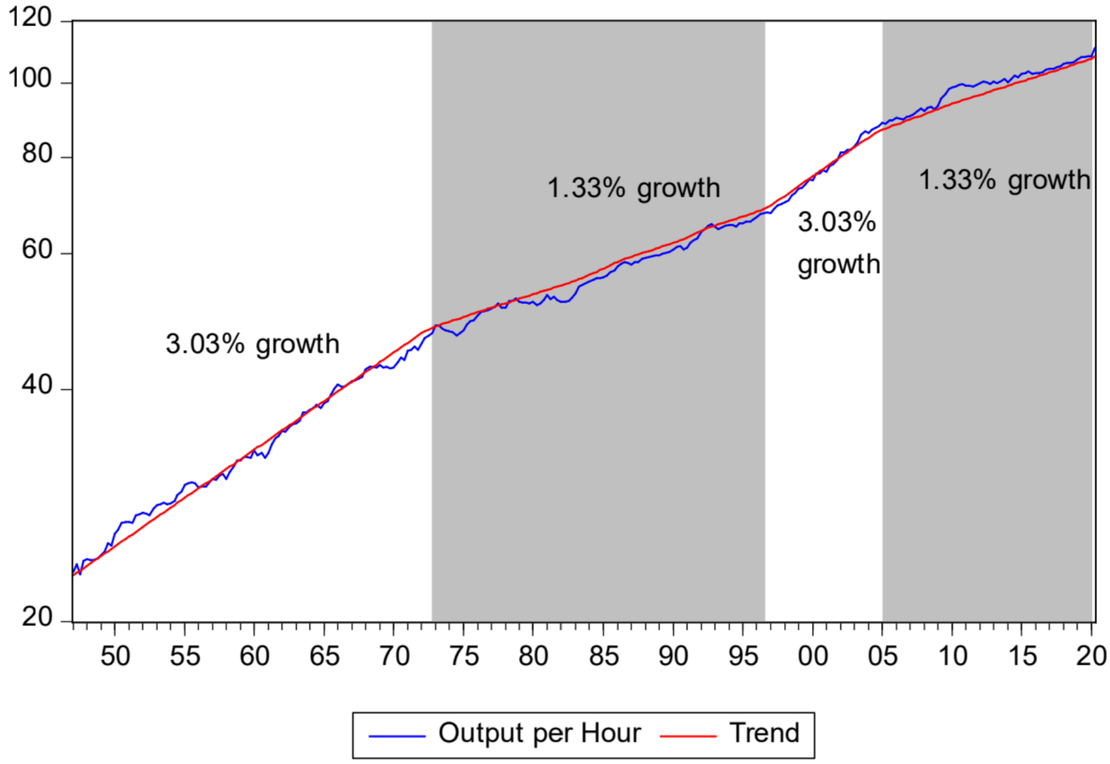


Figure A.1: US labour productivity growth trends (log scale)
Source: Based on Kahn and Rich (2007) and kindly provided by James Kahn.

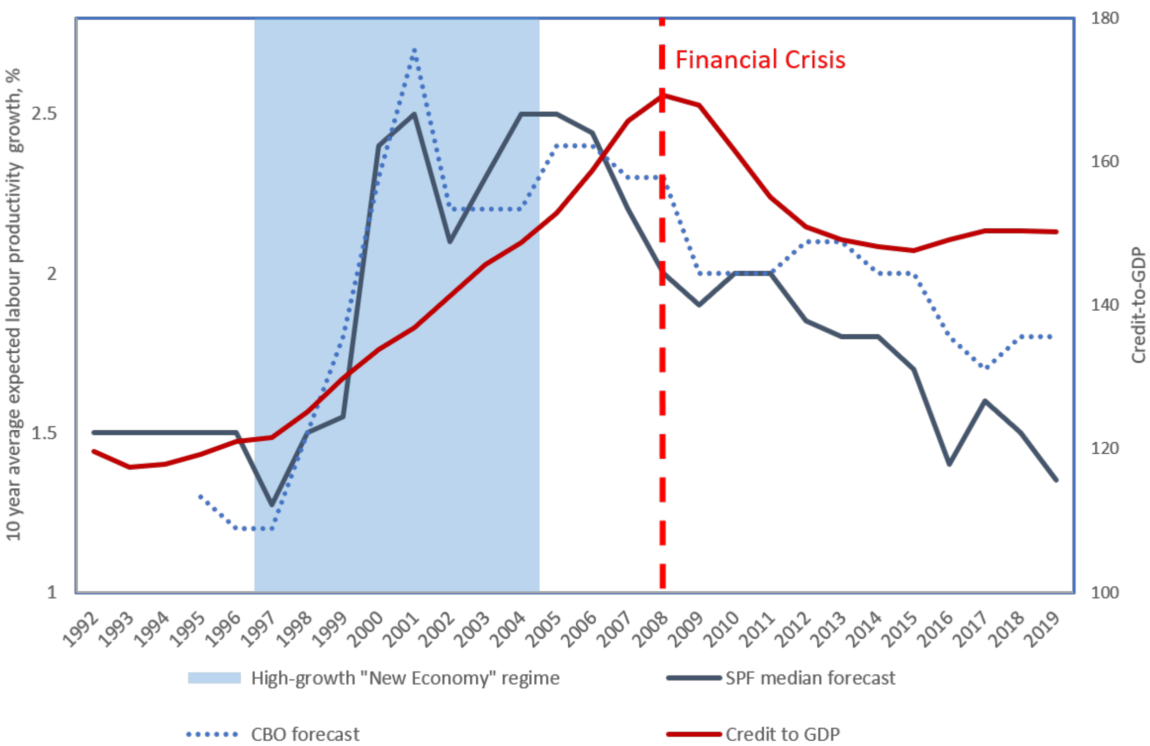


Figure A.2: New Economy, long-run productivity expectations, and credit in the US
Source: Kahn and Rich (2007), BIS, FRB Philadelphia

this period was marked by a large credit boom in the US. As is evident from the figures, however, the trend productivity growth had slowed down again around the beginning of 2005. The evidence that productivity growth had slowed *well before* rather than as a result of the Great Depression is also presented in Fernald (2014); Cette et al. (2016); Cao and L'Huillier (2018). Long-term forecasts began to decline, and the Financial Crisis broke out three years later.

Appendix B. Capital investment

In this appendix I relax the assumption that the stock of capital is fixed.

Suppose that capital is produced by households, who only use their labour as an input. Each household runs a competitive capital producing firm and chooses the level of output to maximize its profit:

$$\max_{k_i} Qk_i - \gamma_i \frac{k_i^2}{2}, \tag{A.1}$$

where k_i is the amount of capital produced, and the quadratic term represents the disutility of labour required to produce k_i units of capital. The first-order condition is

$$k_i(Q) = \frac{Q}{\gamma_i}, \tag{A.2}$$

which, aggregated over all households, yields the aggregate supply of capital:

$$K^S(Q) = \int_{i \in \mathcal{H}} \frac{Q}{\gamma_i} \mathrm{d}F_h(i) = \frac{Q}{\gamma}. \tag{A.3}$$

The following proposition, similarly to Proposition 1 in the paper, establishes that the risk of a financial crisis emerges in equilibrium when the aggregate banks' net worth is sufficiently low, and the probability of a high productivity state is sufficiently high.

Proposition A.1. *There is an essentially unique equilibrium,*[28] *which is parameter-dependent. Suppose the aggregate net worth of the banking sector is sufficiently scarce, namely* $N < \frac{1-\theta}{\gamma}(R^H)^2$. *Then there exist thresholds* $\underline{\pi}, \overline{\pi}$, *with* $\ddot{\pi} < \underline{\pi} < \overline{\pi} < 1$, *such that:*

- *If* $\pi \geq \overline{\pi}$, *all banks choose the risky strategy in equilibrium, and there is a systemic financial crisis if the low state materialises.*
- *If* $\pi \in (\underline{\pi}, \overline{\pi})$, *some banks are safe, and some are risky in equilibrium.*
- *If* $\pi \leq \underline{\pi}$, *all banks choose the safe strategy, and never fail in equilibrium.*

[28] With respect to aggregate outcomes.

Again, $N < \frac{1-\theta}{\gamma} R^{H^2}$ is a plausible assumption. It implies that the banking sector does not have enough own resources to be able to bid up the price of capital to the first-best level even when the high state becomes certain.

The central results therefore remain similar when we allow for the aggregate capital to vary, although it is more challenging to obtain closed-form solutions for the probability thresholds, capital prices, and investment. However, it is still possible to characterize the equilibrium in the limiting case where $N \to 0$, i.e. the banking sector's net worth is very low, and banks are forced to seek most of their financing externally. In this case, the price of a unit of capital approaches the amount of external financing that banks can raise per project, and we have the following result:

Corollary A.1. *As $N \to 0$, both thresholds $\underline{\pi}, \overline{\pi} \to \ddot{\pi}$, and the equilibrium capital price is*

$$Q^* \to \begin{cases} \theta\left(\pi R^H + (1-\pi)R^L\right) & \text{if } \pi \leq \dot{\pi} \\ \pi\,\theta R^L + (1-\pi)R^L & \text{if } \pi \in (\dot{\pi}, \underline{\pi}] \\ \pi\theta R^H + (1-\pi)\Lambda & \text{if } \pi \geq \overline{\pi}. \end{cases} \tag{A.4}$$

The equilibrium capital investment is $K^* = Q^*/\gamma$.

When banks have very low net worth, the ability to raise external financing becomes their crucial concern in equilibrium. Banks are willing to become risky the moment it allows them to better finance themselves, as reflected in $\overline{\pi}$ falling to $\ddot{\pi}$.

Because risky banks are able to raise more financing, thus increasing the demand for capital, we also have:

Corollary A.2. *When π is large and the equilibrium features the risk of a financial crisis, the capital prices and aggregate investment are higher compared to a case in which all banks restrict their leverage to avoid a financial crisis in the L state.*

In other words, the deposit financing indeed enables a boom in asset prices and investment when there is a productivity boom.

Appendix C. Proofs

Proof of Proposition 1. It is straightforward to verify that $\ddot{\pi} < \underline{\pi} < \overline{\pi} < 1$ when $N \in (0, (1-\theta)R^H)$. Let us now prove each part of the proposition consecutively.

a. First, we show that this is indeed an equilibrium. If all banks choose the risky strategy and maximum level of deposits they can repay only in the high state, the aggregate demand for capital is:

$$K^D(Q) = \frac{N}{Q - \pi\theta R^H - (1-\pi)\Lambda},$$

which follows from aggregating (15). Together with the fixed capital supply $K^S = 1$, this immediately implies the equilibrium price of capital in (19). By Lemma 4, banks are indeed behaving optimally iff $Q^* \leq \overline{Q}$, which, with a little bit of algebra, can be shown to hold iff $\pi \geq \overline{\pi}$, as required.

To prove uniqueness, suppose there is a different equilibrium in which fraction $\zeta < 1$ of banks (measured by net worth) are risky, and the rest $1 - \zeta$ are safe. Aggregate demand for capital is then given by

$$K^D = \frac{\zeta N}{Q - \pi\theta R^H - (1-\pi)\Lambda} + \frac{(1-\zeta)N}{Q - \pi\theta R^L - (1-\pi)R^L}. \tag{A.5}$$

Combined with the equilibrium condition $K^D = K^S \equiv 1$, it is straightforward to show that the equilibrating price Q is an increasing function of ζ. Therefore, in this other equilibrium the price of capital is below Q^* found in (19), and thus also below $\overline{Q}$. But then all banks would strictly prefer the risky strategy by Lemma 4, a contradiction.

b. To show that it is indeed an equilibrium outcome, first note that because the equilibrium price is $\overline{Q}$, individual banks are indifferent between the safe and risky strategies. Combining (18) with the aggregate capital demand in (A.5) and the market clearing condition, it is straightforward to show that $\overline{Q}$ is indeed the equilibrating price when exactly fraction ζ^* of banks (measured by net worth) choose the risky level of deposits.

   Formally, there are multiple equilibria in this case, since there are numerous strategy profiles, pure and mixed, that result in fraction ζ^* of banks choosing the risky level of deposits. The aggregate equilibrium outcome described in the proposition, however, is unique. There is no equilibrium in which more than fraction ζ^* of banks chose the risky level of deposits. That would lead to more financing and a higher capital price in period 0, but then all banks would strictly prefer to be safe. Similarly, there is also no equilibrium in which less than fraction ζ^* of banks end up choosing the risky strategy, because that would imply a lower capital price and all banks preferring the risky strategy.

c. When $\pi \leq \underline{\pi}$, there is no equilibrium in which any of the banks end up being risky. Firstly, when $\pi \leq \ddot{\pi}$, all banks are safe by Lemma 4. Secondly, if such an equilibrium where to exist when $\pi \in (\ddot{\pi}, \underline{\pi}]$, the demand for capital would be given, again, by (A.5). But then it is straightforward to verify that the equilibrating capital price would always exceed $\overline{Q}$ whenever $\zeta > 0$. But then all banks would strictly prefer being safe, a contradiction.

Q.E.D.

Proof of Proposition 2. We are going to proceed with the proof as follows. First, we derive the condition (20). This happens to be the condition under which the regulator prefers to supervise banks when the probability of the high state is exactly $\underline{\pi}$. Second, we demonstrate that if the condition (20) is satisfied, there exist π_1, π_2, with $\pi_1 < \underline{\pi} < \pi_2 < 1$, such that banks are supervised iff $\pi \in (\pi_1, \pi_2)$. Lastly, we show that it is never optimal for the regulator to supervise banks if it is not optimal to do so when $\pi = \underline{\pi}$, i.e. (20) does not hold.

Without bank supervision, the equilibrium price of capital is

$$Q^*(\pi) = \begin{cases} \min\Big\{N + \theta[\pi R^H + (1-\pi)R^L],\ \pi R^H + (1-\pi)R^L\Big\} & \text{if } \pi \le \dot{\pi} \\ \min\Big\{N + \pi\theta R^L + (1-\pi)R^L,\ \pi R^H + (1-\pi)R^L\Big\} & \text{if } \pi \in (\dot{\pi}, \underline{\pi}] \\ \overline{Q}(\pi) & \text{if } \pi \in (\underline{\pi}, \overline{\pi}) \\ N + \pi\theta R^H + (1-\pi)\Lambda & \text{if } \pi > \overline{\pi}, \end{cases} \tag{A.6}$$

where the last two cases are already given in Proposition 1. The first two cases follow from either: (1) aggregating capital demand from Lemma 2 and combining it with the market clearing condition (under the assumption that banks are constrained and exhaust their financing capacity), or (2) the maximum price that unconstrained banks would pay for capital, $\mathbb{E}[R^S] = \pi R^H + (1-\pi)R^L$, whichever is *smaller*. It is straightforward to verify that $Q^*(\pi)$ is continuous on $[0,1]$.

With bank supervision, the price of capital is simply

$$Q_p(\pi) = \min\Big\{N + \theta\chi[\pi R^H + (1-\pi)R^L],\ \chi[\pi R^H + (1-\pi)R^L]\Big\}. \tag{A.7}$$

The first term in the min operator again follows from the capital market clearing condition assuming that banks are constrained and exhaust their financing capacity. The second term caps the equilibrating price at the maximum level that unconstrained banks would pay for capital under supervision, $\chi\mathbb{E}[R^S]$. $Q_p(\pi)$ is an increasing and continuous function on $[0,1]$.

At $\pi = \underline{\pi}$, banks are constrained and so $Q^*(\underline{\pi}) = N + \underline{\pi}\theta R^L + (1-\underline{\pi})R^L$. The household-friendly regulator will find it optimal to supervise banks if and only if $Q_p(\underline{\pi}) > Q^*(\underline{\pi})$, from which the condition (20) follows.

Suppose the condition (20) is satisfied, and $Q_p(\underline{\pi}) > Q^*(\underline{\pi})$. Then there is $\pi_1 \in (\dot{\pi}, \underline{\pi})$ s.t. $Q_p(\pi) > Q^*(\pi)$ when $\pi \in (\pi_1, \underline{\pi}]$, but $Q_p(\underline{\pi}) \le Q^*(\pi)$ when $\pi \le \pi_1$. This follows from first recognizing that we would always have $Q^*(\pi) > Q_p(\pi)$ when $\pi \le \dot{\pi}$, and also when banks are unconstrained and $Q^* = \mathbb{E}[R^s]$. Therefore, on the interval $\pi \in (\dot{\pi}, \underline{\pi})$, we can have $Q_p(\pi) > Q^*(\pi)$ only when banks are constrained and $Q^*(\pi) = N + \pi\theta R^L + (1-\pi)R^L$. This is a decreasing function of π, and so must intersect $Q_p(\pi)$ at a single point, π_1.

There is also $\pi_2 \in (\underline{\pi}, 1)$ s.t. $Q_p(\pi) > Q^*(\pi)$ when $\pi \in [\underline{\pi}, \pi_2)$, but $Q_p(\pi) \le Q^*(\pi)$

when $\pi \geq \pi_2$. This follows because, firstly, $Q_p(\pi)$ must cross $Q^*(\pi)$ at least once, since it is easy to verify that $Q^*(1) > Q_p(1)$. Secondly, once $Q_p \leq Q^*$ at some $\hat{\pi} > \underline{\pi}$, then $Q_p(\pi) \leq Q^*(\pi)\ \forall \pi \in [\hat{\pi}, 1]$. Let us now show the latter.

First, it is straightforward to see that if $Q_p \leq Q^*$ at some $\hat{\pi} \in [\underline{\pi}, \overline{\pi}]$, then $Q_p(\pi) \leq Q^*(\pi)\ \forall \pi \in [\hat{\pi}, \overline{\pi}]$. This is because on the interval $[\underline{\pi}, \overline{\pi}]$, $\overline{Q}(\pi)$ is a linear function with a slope greater than $R^H - \Lambda$, and so always greater than the slope of $Q_p(\pi)$, which is at most $\chi(R^H - R^L)$. Next, we also have that if $Q_p \leq Q^*$ at $\hat{\pi} \in [\overline{\pi}, 1]$, then $Q_p(\pi) \leq Q^*(\pi)\ \forall \pi \in [\hat{\pi}, 1]$. This is straightforward to see if, when supervised, banks are constrained in equilibrium at $\hat{\pi}$, and so $Q_p(\pi) = N + \theta\chi[\pi R^H + (1-\pi)R^L]$ on $[\hat{\pi}, 1]$, which again has a lower slope than that of $Q^*(\pi)$. If, when supervised, banks are unconstrained at $\hat{\pi}$, then $Q_p(\hat{\pi}) = \chi[\hat{\pi} R^H + (1-\hat{\pi})R^L]$. With a bit of algebra, we can verify, however, that $Q_p(\hat{\pi}) \leq Q^*(\hat{\pi})$ at this point also implies that $N + \theta\chi[\hat{\pi} R^H + (1-\hat{\pi})R^L] < Q^*(\hat{\pi})$.[29] This, in turn, implies that Q_p cannot possibly exceed Q^* as π increases.

Taken together, the above implies that the regulator supervises banks iff $\pi \in (\pi_1, \pi_2)$ when (20) holds.

The above considerations about the behaviour of the functions $Q^*(\pi)$ and $Q_p(\pi)$ also lead us to conclude that there is no π at which the regulator would find it optimal to supervise banks when the condition (20) is not satisfied. As $Q_p \leq Q^*$ at $\underline{\pi}$, it must be the case that $Q_p \leq Q^*$ when $\pi < \underline{\pi}$, since even in the region where banks are constrained and $Q^*(\pi) = N + \pi\theta R^L + (1-\pi)R^L$, $Q^*(\pi)$ increases and $Q_p(\pi)$ falls when π decreases below $\underline{\pi}$. Similarly, $Q_p \leq Q^*$ at $\underline{\pi}$ implies that $Q_p(\pi) \leq Q^*(\pi)\ \forall \pi > \underline{\pi}$. *Q.E.D.*

[29] Using $Q_p(\hat{\pi}) \leq Q^*(\hat{\pi})$, it follows that $N + \theta\chi[\hat{\pi} R^H + (1-\hat{\pi})R^L] < Q^*(\hat{\pi})$ whenever $\theta N < (1-\theta)[\hat{\pi}\theta R^H + (1-\hat{\pi})\Lambda]$. The right-hand-side of the last inequality increases in $\hat{\pi}$. This inequality holds, however, even when $\hat{\pi} = \overline{\pi}$, which can be verified by substituting in $\overline{\pi}$ from Proposition 1 and rearranging.