A Macroeconomic Model of Endogenous Systemic Risk Taking*

David Martinez-Miera
Universidad Carlos III

Javier Suarez
CEMFI and CEPR

First version: December 2010
This version: June 2011
Preliminary and incomplete

Abstract

We analyze systemic risk taking in a simple dynamic general equilibrium model with banks. Banks collect funds from final savers and make loans to firms, which use them to pay for capital and labor in advance. Banks are owned and managed by risk-neutral bankers who provide the equity needed to comply with capital requirements. Bankers reinvest their wealth in banks' equity insofar as the endogenous scarcity rents provide sufficient incentives for doing so. Bankers decide their (unobservable) exposure to systemic shocks by trading off risk-shifting gains with the value of preserving their capital after a systemic shock. The analysis identifies dynamic trade-offs relevant for the optimal design of capital requirements, which reduce banks' systemic risk taking (and hence the losses caused by systemic shocks) but also reduce credit and output in *normal times*. Optimal capital requirements under our calibration of the model are quite high (14%) despite their negative impact on GDP is much bigger than commonly estimated.

Keywords: Capital requirements, Risk shifting, Credit cycles, Systemic risk, Financial crises, Macroprudential policies.

JEL Classification: G21, G28, E44

* We would like to thank Matthieu Darracq-Paries, Martin Ellison, Carlos González-Aguado, Hans Gersbach, Claudio Michelacci, Diego Rodriguez-Palenzuela, and seminar audiences at Bank of Japan, National Bank of Belgium, Swiss National Bank, and Universidad de Navarra for helpful comments and suggestions. We acknowledge financial support from the European Central Bank and from the Spanish Ministry of Science and Innovation through the Consolider-Ingenio 2010 Project "Consolidating Economics." Address for correspondence: CEMFI, Casado del Alisal 5, 28014 Madrid, Spain. Phone: +34-914290551. E-mail: suarez@cemfi.es.

1 Introduction

This paper analyzes the role of banks in generating endogenous systemic risk. We consider a canonical problem of excessive risk-taking by banks in a dynamic general equilibrium model in which banks are subject to capital requirements and the supply of equity funding to banks is limited by the wealth endogenously accumulated by the bank insiders (the *bankers*). We find that banks' systemic risk-taking can be generally reduced by increasing bank capital requirements. Yet doing so has negative implications for the levels of credit and output, producing static and dynamic trade-offs relevant for the socially optimal design of capital requirements.

We consider an economy subject to *systemic shocks* which, like in Rancière at al. (2008), are small probability events in which certain investments fail in a highly correlated manner, provoking the default of the loans financing them.[1] Importantly, banks make an explicit but *unobservable* decision on their degree of exposure to these shocks—or, equivalently, on the extent to which the loans in their portfolio tend to default in a correlated manner.[2] Banks' temptation to undertake highly correlated investments is due to the risk-shifting incentives of levered firms, like in Jensen and Meckling (1976) and many other corporate finance and banking models.[3]

In our model, banks collect deposits from some final savers and are the exclusive providers of loans to perfectly competitive firms which need them to pay for physical capital and labor in advance. Firms' production processes are subject to failure risk and, depending on their degree of exposure to systemic shocks, may be systemic or non-systemic. For a given combination of physical capital and labor and conditional on their success or failure, all pro-

[1] Systemic shocks resemble the *rare economic disasters* considered in Rietz (1988) and Barro (2009), among others. We treat systemic shocks as exogenous and focus on the endogenous exposure of banks to them to avoid the complexity associated with modelling the mechanisms that generate correlation (bubbles, negative spillovers caused by fire sales and disordered liquidation, interbank linkages, and informational issues that support the extension of panic among bank investors). A systemic shock can empirically correspond to phenomena such as the bust of the US housing market (and its implications for subprime mortgages, securitization markets, and money markets) around the Summer of 2007.

[2] A similar correlation choice has been modeled and analyzed in the microeconomic banking literature by Acharya and Yorulmazer (2007).

[3] See Section 2 for additional references.

duction processes yield the same output. However, non-systemic production processes have stochastically independent, identically distributed failures. In contrast, systemic production processes fail simultaneously if the negative systemic shock occurs (and in an independent fashion if it does not).

Following the literature on risk shifting, we assume that systemic firms have probabilities of failure that, conditional on not suffering the systemic shock, are lower, but unconditionally are higher, than those of non-systemic firms. Hence, systemic firms generate lower expected net present value than non-systemic firms but may be a tempting investment opportunity for highly levered institutions such as banks. We assume that the unobservable systemic or non-systemic nature of each firm's production process is a joint decision made by the firm and its bank when establishing the terms of their lending relationship.

Realistically and in line with many models in the literature on bank risk-taking, we assume that (i) bank deposits are fully insured by the government and (ii) banks are subject to capital requirements. We model the dynamics of bank capital along the same lines as Gertler and Karadi (2009), Gertler and Kiyotaki (2009), and Meh and Moran (2010). This implies thinking of bank capital as the funds provided by the special class of agents who own and manage the bank—the *bankers*. The maximal available amount of this funding is determined by bankers' wealth, whose endogenous dynamics is affected by the profits and losses made by the banks in each period.[4]

Bankers may be interested in making systemic loans because they are protected by limited liability and, if their bank fails, part of the losses go to the government as the provider of deposit insurance.[5] However, bank capital requirements influence bankers' incentives in regards to the adoption of systemic risk. They do so through the conventional leverage-reduction effect, that reduces the static gains from risk-taking, and also by increasing the scarcity of bank capital in each state of the economy, which is important from a dynamic optimization perspective.

[4]Gertler et al. (2010) consider a setup where bankers' *inside equity* can be complemented with *outside equity* but an agency problem limits the access to the latter to a certain multiple of the former.

[5]A simillar effect would occur vis-a-vis bank depositors if the deposits were uninsured. What is really important for the distortion is limited liability and the unobservability of the bank's risk-taking decision.

Indeed, the anticipated value of a unit of bank capital in the future is key to bankers' systemic risk-taking decisions. When a banker's capital is devoted to make systemic loans, the banker obtains higher returns (and accumulates more capital) insofar as the systemic shock does not occur. But if the systemic shock occurs, the invested equity is lost. In contrast, if the banker invests in non-systemic loans, he receives a lower return in *normal times* (i.e. if the systemic shock does not realize), but preserves his capital when the systemic shock occurs. Hence, the destruction of bank capital allows surviving capital to earn higher scarcity rents after a crisis and produces a "last bank standing effect" like in Perotti and Suarez (2002). This effect provides an incentive for bankers not to be so inclined towards systemic lending. Of course, it is crucial for this mechanism to operate in practice that systemic crises are resolved with a maximal dilution of failed bankers' equity stakes in their banks, even if in reality, due to considerations that the model does not capture, the banks as organizations were bailed-out and survived the crisis.

Systemic risk taking has negative static and dynamic implications in our model. First, even from a single-period investment horizon perspective, systemic firms generate less overall expected net present value than non-systemic firms. Second, when the systemic shock realizes, the economy suffers a loss of aggregate bank capital which in turn produces a credit crunch, and some corresponding output and net consumption losses during the transition periods it takes for the economy to recover its pre-crisis levels of bank capital and output.

We show that strengthening capital requirements reduces bankers' systemic risk taking through the two channels mentioned above, thus making the economy less exposed to the inefficiency of systemic investments and the contractions that follow the realization of systemic shocks. However, the gains on the risk-taking front come at the cost of reducing credit, output, and wages (especially in the *pseudo-steady state* reached after sufficiently many periods without experiencing the systemic shock), producing non-trivial welfare trade-offs. Measuring welfare as the expected present value of total net consumption flows in the economy, we show that these trade-offs generally produce a unique interior social welfare maximizing level of capital requirements.

Under our tentative calibration, social welfare is indeed maximized under positive and relatively large (flat) capital requirements. Specifically, it is optimal to require banks to finance 14% of their loans with bankers' wealth. To fix ideas, we compare the performance of the economy under such minimum capital requirement with an alternative scenario with a (low) 7% capital requirement. We find that moving from the low to the optimal capital requirement reduces the fraction of systemic loans in the pseudo-steady state from 72% to 25% and increases our overall measure of social welfare in an amount equivalent to a perpetual increase of 0.9% in aggregate net consumption.

Importantly, moving capital requirements from 7% to 14% brings about apparent losses in terms of common macroeconomic aggregates such as bank credit, GDP, and wages, whose pseudo-steady state values fall on average by 22%, 7%, and 9%, respectively. However, the pseudo-steady state values of these variables are bad proxies of social welfare in the presence of systemic risk since they neglect important losses that only become apparent when the risk materializes.[6] Our analysis also calls into question the use of GDP (and especially its level in normal times) as a summary statistic of the macroeconomic effects of macroprudential policies.

The rest of the paper is organized as follows. Section 2 places the contribution of the paper in the context of the existing literature. Section 3 describes the model. Section 4 derives the conditions relevant for the definition of equilibrium. Section 5 describes our calibration exercise and the main quantitative results. Section 6 discusses a number of extensions. Section 7 concludes. The Appendix describes the numerical method used to solve for equilibrium and provides several equivalent expressions for social welfare.

2 Related literature

Our paper is related to recent efforts to incorporate banks and their (endogenous) contribution to systemic risk into core macroeconomic analysis. Dynamic stochastic general

[6] For instance, the fall in aggregate net consumption in the year that follows a systemic shock is (relative to the pseudo-steady state level) of 11.7% with capital requirements of 7% and of only 2.8% with capital requirements of 14%. Similarly, the falls in bank credit, GDP, and wages in the year after a shock are of 65%, 30%, and 37% with capital requirements of 7%, while it is of 24%, 9%, and 11% with capital requirements of 14%.

equilibrium (DSGE) models in use by central banks prior to the beginning of the crisis (e.g. in the tradition of Smets and Wouters, 2007) paid no or very limited attention to financial frictions. Several models considered idiosyncratic default risk and endogenous credit spreads using the framework provided by Bernanke et al. (1999) but very few were explicit about banks. Van den Heuvel (2008) introduces banks' liquidity provision function in a macroeconomic setup in order to assess the welfare costs of capital requirements which in his model are introduced as an ad hoc piece of regulation with no relationship to risk-taking incentives. Meh and Moran (2010) look at the type of market-imposed capital requirements of Holmström and Tirole (1997), i.e. those that emerge in a setup in which banks' outside financiers are not protected by government guarantees and bankers make costly unobservable decisions regarding the effectiveness of their monitoring.

More recently, various authors—including Agénor et al. (2009), Christiano et al. (2010), and Darracq-Pariès et al. (2010)—have extended models in the DSGE tradition with the explicit goal of capturing banking frictions. However, the reduced-form approach typically leaves aside an explicitly microfounded role for the introduced regulatory ingredients and impedes a fully-fledged welfare analysis.

Because of our modeling of the dynamics of bank capital and the existence of an explicit connection between bank capital and bankers' incentives, the papers most closely related to ours are Gertler and Karadi (2009), Gertler and Kiyotaki (2009), and the already-commented Meh and Moran (2010). These papers prescribe for bankers' wealth the same type of dynamics previously postulated for entrepreneurial net worth in the models with financial constraints, e.g., Carlstrom and Fuerst (1997), and Kiyotaki and Moore (1997). In Gertler and Karadi (2009) and Gertler and Kiyotaki (2009), requiring bankers to fund some minimal fraction of their banks with their own funds prevents a potential fund diversion problem, like in Hart and Moore (1994).

Similar bank capital dynamics is postulated by Brunnermeier and Sannikov (2009), who also model crises as the result of large but infrequent shocks. However, they put the emphasis on identifying channels through which fire sales are a source of pecuniary externalities. In their model, the interaction of asset price volatility interacts with value-at-risk based capital requirements plays a key role in the transmission of the shocks, but optimal capital regulation is not discussed.[7]

Our focus on explicit risk-taking decisions and optimal capital regulation connects our contribution to long traditions in the corporate finance and banking literatures. Reviewing the many papers that have dealt with the excessive risk taking incentives of levered firms and banks in partial equilibrium (and frequently static) setups exceeds the scope of this section. We will refer here to just a few selected pieces, referring the interest reader to Bhattacharya, Boot and Thakor (1998) and Freixas and Rochet (2008) for excellent reviews.

Earlier references on risk-shifting include Jensen and Meckling (1976), who consider the excessive risk-taking of levered firms in a corporate finance context, and Stiglitz and Weiss (1981), who analyze borrowers' risk-taking incentives in a credit market equilibrium context. Excessive risk-taking by banks is already alluded by Kareken and Wallace (1978) as an important side effect of deposit insurance and a motivation for regulation.[8] The role of capital requirements in ameliorating this problem and their interaction with the dynamic incentives due to banks' franchise values is a central theme in Hellmman, Murdock and Stiglitz (2000) and Repullo (2004), where banks earn rents due to market power. The dynamic incentives for prudence associated with the rise in the franchise value of the banks that survive a systemic crisis were first pointed out by Perotti and Suarez (2002), and are also central to the analysis of Acharya and Yorulmazer (2007, 2008). Differently from all these papers, the banks in our model are perfectly competitive but bank capital earns scarcity rents due to the fact that bankers are the sole providers of bank capital and their endogenously accumulated wealth is limited. Thus, the shadow value of bank capital plays an incentive role similar to the role assigned to banks' franchise value in the prior literature.

Finally, our paper abstracts, for simplicity, from the entrepreneurial-incentives channel

[7] They focus on the polar case in which capital requirements guarantee that banks never fail.

[8] When some relevant dimension of risk taking is unobservable (as it is the case in our model), equilibrium risk-taking may be excessive even without government guarantees on bank liabilities. Of course, the underpricing of those guarantees (or their flat pricing in setups where some form of risk-sensitivity were feasible) tends to worsen the problem. Dewatripont and Tirole (1994) describe safety net guarantees as part of a social contract whereby banks' unsophisticated or unmotivated-to-monitor creditors delegate the task of controlling banks' risk taking on the supervisory authorities who, in exchange, provide explicit or implicit insurance to their debts.

emphasized by Boyd and De Nicoló (2005) in their reexamination of the (traditionally accepted as negative) link between market power and bank solvency. In essence, their story goes, higher loan rates may damage entrepreneurial incentives and have an indirect negative impact, via larger loan defaults, on banks' solvency.[9] Extending our framework for the explicit consideration of entrepreneurial incentives would be feasible, e.g. under a formulation similar to Holmström and Tirole (1997) or Repullo and Suarez (2000), but given the length of the current paper should rather be left for future research.[10]

3 The model

Consider a perfect competition, infinite horizon model in discrete time $t = 0, 1, ...$

3.1 Agents

There are two wide classes of risk-neutral agents in the economy: *patient* agents, who essentially act as providers of funding to the rest of the economy, and *impatient* agents, who include pure workers, bankers, and entrepreneurs. Additionally, there is a *government* which provides deposit insurance and decides on bank capital requirements and any other element of macroprudential policy.

Patient agents have deep pockets. Their required expected rate of return (which can be interpreted as the exogenous return on some risk-free technology) is r per period. They are the natural financiers of the remaining agents but we assume that, due to unmodeled informational and agency frictions, they cannot directly lend to the final borrowers. They provide a perfectly elastic supply funds to *banks* in the form of deposits.[11]

Impatient agents, of whom there is a continuum of measure one, are infinitely lived,

[9] Martinez-Miera and Repullo (2010) qualify this result in the presence of non-perfectly correlated loan defaults and conclude that the link between bank competition and the risk of bank insolvency is generally U-shaped.

[10] In such an extension, it would be natural to explicitly allow for the endogenous accumulation of entrepreneurial net worth which, under the simplest possible modeling, would become an additionally aggregate state variable and, thus, make the analysis computationally more demanding.

[11] In an open economy interpretation, one can think of patient agents as international capital market investors and r as the international risk-free rate.

have a discount factor $\beta < 1/(1+r)$, and inelastically supply a unit of labor per period at the prevailing wage rate w_t. Most impatient agents are just *workers*, but some workers temporarily become *bankers*—i.e. receive some especial skills or qualifications needed to own and manage a bank—and some temporarily become *entrepreneurs*—i.e. acquire the ability to create and run a firm. Specifically, each worker has a small independent probability $\phi\psi$ of learning in each date t that he will become an active *banker* at date $t+1$. In parallel, each banker active at date t has a small independent probability ψ of losing his skills or qualifications by the end of the period and becoming just a worker again at date $t+1$. This produces a stationary size ϕ for the (changing) population of active bankers. Finally, a tiny fraction μ of the patient agents who at each date do not have the opportunity to act as bankers receive the opportunity of acting as entrepreneurs (i.e. owning and managing a firm) for just one period.

Below we will adopt the assumptions needed to guarantee that patient agents take the opportunity to act as bankers or entrepreneurs whenever available.[12] We also assume that the probabilities $\phi\psi$ and μ are small enough (relative to the value of equity devoted to banking and entrepreneurial activities) for the accumulation of wealth by workers not to be worthy prior to learning about their conversion into bankers or entrepreneurs. Together with our assumption that entrepreneurs are only active for one period, this allows us to focus on the accumulation of wealth by bankers and would-be bankers (workers who learn that they will be bankers in the next period).

3.2 Firms

The measure μ continuum of entrepreneurs active in every period run an equal measure of perfectly competitive firms indexed by $i \in [0, \mu]$. Each firm operates a constant returns to scale technology that allows it to transform the capital k_{it} and labor n_{it} employed at t into

$$y_{it+1} = z_{it+1}[AF(k_{it}, n_{it}) + (1-\delta)k_{it}] + (1 - z_{it+1})(1-\lambda)k_{it} \tag{1}$$

[12] In the current version of the paper, entrepreneurs play a pretty passive role and eventually receive a competitive profit of zero at all dates.

units of the consumption good (which is the numeraire) at $t+1$. The term $z_{it+1} \in \{0,1\}$ is a binary random variable realized at $t+1$ that indicates whether the firm's production process succeeds or fails, the parameters δ and $\lambda \geq \delta$ are the rates at which capital depreciates when the firm succeeds and when it fails, respectively, A is total factor productivity, and

$$F(k_i, n_i) = k_i^{\alpha} n_i^{1-\alpha}, \tag{2}$$

with $\alpha \in (0,1)$.

Thus firms fail to produce output on top of depreciated capital whenever $z_{it+1} = 0$. To capture different degrees of exposure to a common *systemic* shock $\varepsilon_{t+1} \in \{0,1\}$, we assume that the technology can be operated in two modes: non-systemic ($\xi_{it} = 0$) and systemic ($\xi_{it} = 1$). For firms operating under the non-systemic mode, z_{it+1} is independently and identically distributed across firms with

$$\Pr[z_{it+1} = 0 \mid \varepsilon_{t+1} = 0,\, \xi_{it} = 0] = \Pr[z_{it+1} = 0 \mid \varepsilon_{t+1} = 1,\, \xi_{it} = 0] = p_0.$$

In contrast, all firms operating under the systemic mode have

$$\Pr[z_{it+1} = 0 \mid \varepsilon_{t+1} = 0,\, \xi_{it} = 1\,] = p_1 < \Pr[z_{it+1} = 0 \mid \varepsilon_{t+1} = 1,\, \xi_{it} = 1] = 1,$$

where failure in case of no shock ($\varepsilon_{t+1} = 0$) is independently distributed across firms. Hence, systemic firms fail simultaneously if the negative systemic shock occurs ($\varepsilon_{t+1} = 1$), and independently, with probability $p_1 < 1$, otherwise.

The failure rate associated to any positive measure of non-systemic firms is, by the law of large numbers, constant and equal to p_0, while the failure rate among a positive measure of firms operated in the systemic mode will be:

$$z_{t+1} = \begin{cases} 1 & \text{if } \varepsilon_{t+1} = 1, \\ p_1 & \text{otherwise.} \end{cases} \tag{3}$$

We assume that the negative systemic shock, $\varepsilon_{t+1} = 1$, occurs with a constant independent small probability ε at the end of each period. Finally, as in the *risk-shifting* literature, we assume:

(i) $p_1 < p_0$, which will make the lending to systemic firms attractive to bankers, and

(ii) $E(z_{it+1} \mid \xi_{it} = 1) = (1-\varepsilon)p_1 + \varepsilon > E(z_{it+1} \mid \xi_{it} = 0) = p_0$, which implies that systemic firms are overall less efficient than non-systemic ones and makes risk-shifting socially undesirable.

All entrepreneurs running firms at date t are penniless, are protected by limited liability, and maximize the expected payoffs obtained from their firms at $t+1$, point at which they become just workers again. Limited liability, which may be interpreted as an exogenous institutional constraint or an implication of anonymity, implies that entrepreneurs' contemporaneous or future wages cannot be used as collateral for entrepreneurial activities. This allows us to abstract from wealth accumulation by entrepreneurs.

Firms' activities are financed with bank loans. Specifically, as in Bruche and Suarez (2010), we assume that each firm requires a bank loan of size $l_{it} = k_{it} + w_t n_{it}$ to pay "in advance" for the capital k_{it} and labor n_{it} utilized at date t. The loan implies the *promise to repay* $B_{it} \leq AF(k_{it}, n_{it}) + (1-\delta)k_{it}$ at $t+1$. As in a standard debt contract, this will involve an effective repayment B_{it} if the firm does not fail and $\min\{B_{it}, (1-\lambda)k_{it}\} = (1-\lambda)k_{it}$ if the firm fails.[13]

The variables in the tuple $(\xi_{it}, k_{it}, n_{it}, l_{it}, B_{it})$ are set by mutual agreement between each firm and its bank at date t. Importantly, as we further discuss below, a firm's systemic orientation is private information of the firm and its bank, ruling out regulations directly contingent on ξ_{it}.

3.3 Banks

Regulation, as specified below, obliges banks to finance part of their activity with equity capital, i.e. with funds coming from bankers' accumulated wealth. Banks combine these funds with funds collected as fully-insured one-period deposits among the patient agents (and perhaps the impatient agents who save their labor income in order to be able to invest it as bank capital in the next date) and devote all funds to make one-period loans to firms.[14]

[13] Notice that with non-negative loan rates and wages, we will necessarily have $B_{it} \geq l_{it} = k_{it} + w_t n_{it} \geq k_{it} \geq (1-\lambda)k_{it}$.

[14] We assume that impatient agents cannot borrow from banks for consumption purposes. This could be justified as a result of the impossibility of pledging their future income due to, e.g., intertemporal anonymity.

Prudential bank regulation obliges banks to hold *granular* loan portfolios, that is, to extend infinitesimal loans to a continuum of firms. This guarantees that banks fully diversify away firms' idiosyncratic failure risk. In contrast, diversification does not eliminate the systemic risk associated with the lending to systemic firms.

Because of convexities induced by limited liability, bankers will find it optimal to specialize their banks in either non-systemic or systemic loans.[15] Banks are assumed to be perfectly competitive and operate under constant returns to scale, so we can refer w.l.o.g. to a representative *non-systemic bank* ($\xi = 0$) and a representative *systemic bank* ($\xi = 1$). We will denote by $e_{\xi t}$ the amount of bankers' wealth invested in bank $\xi = 0, 1$ at date t. The deposits taken and the loans made by each bank at date t will be denoted $d_{\xi t}$ and $l_{\xi t}$, respectively. Each bank's balance sheet constraint imposes

$$l_{\xi t} = d_{\xi t} + e_{\xi t}, \tag{4}$$

for $\xi = 0, 1$.[16]

The allocation of bank capital to each bank takes place in a perfectly competitive fashion. Because of limited liability, bankers contributing $e_{\xi t}$ to bank ξ at date t will receive the free cash flow of the bank at $t+1$ (i.e. repayments from loans net of repayments due to deposits) if it is positive and zero otherwise. Bankers optimally decide the allocation of their wealth based on their expectation about bank equity returns (and the valuation of the resulting wealth) across different possible states at $t+1$, which they take as given. In principle, bankers can invest their wealth as capital of the non-systemic bank, capital of the systemic bank or insured deposits, and they can also decide to consume all or part of their wealth at t.

At the other side of the market, banks take as given bankers' value-weighted required return on wealth (and their valuation of wealth across possible states at $t+1$) to formulate

Consistently with this view, we might argue that banks can borrow from the patient agents (taking one-period deposits) and firms from the banks (using one-period loans) because of the *pledgeability* of the assets that they own at the end of each period.

[15] For a formal argument, see Repullo and Suarez (2004).

[16] Given that both classes of banks have access to unlimited funding at the risk-free rate (via deposits from the patient agents), we can abstract from the analysis of interbank lending and borrowing.

the participation constraint which guarantees that bankers are willing to provide the equity funding $e_{\xi t}$ that each bank needs at t. This participation constraint is taken into account when setting the terms of the lending contracts $(\xi_{it}, k_{it}, n_{it}, l_{it}, B_{it})$ with each of the entrepreneurs that they finance.

Summing up, in our economy, bankers are the agents who solve the genuinely dynamic optimization (portfolio) problems that determine the part of bankers' wealth eventually invested in equity of the non-systemic bank e_{0t} or in equity of the systemic bank e_{1t}. Banks, instead, are perfectly competitive one-period enterprises in which bankers invest (and which are run in bankers' interest). For future use, we will denote by $x_t \equiv e_{1t}/e_t \in [0, 1]$ the fraction of total bank capital e_t invested in systemic banks.

3.4 The government

The systemic orientation of each bank is supposed to remain private information of its owners, impeding the enforcement of regulations directly contingent on it.[17] We further assume that banks play a pooling equilibrium in which the non-systemic bank makes individually optimal decisions when contracting with firms, while the systemic bank prevents being identified as such by mimicking such behavior in every aspect but the unobservable systemic orientation $\xi_{it} = 1$ of the firms receiving its loans. Since a systemic bank might additionally be detected by observing the value of its realized return on equity at $t+1$ (namely, R_{1t+1}, which will tend to be higher than the return on equity of the non-systemic bank, R_{0t+1}, if the systemic shock does not occur), we also assume that sufficient lack of transparency in bank accounting and managerial compensation practices allow the owners of systemic banks to appropriate the excess normal-time returns without being discovered.

For simplicity, we assume that the government provides full insurance (principal and interest) to deposit claims in case of bank failure. The cost of this scheme is assumed to be covered on a pay-as-you go basis with the revenue raised by some non-distortionary tax on the impatient agents (e.g. on the consumption of the workers), which prevents the possibility

[17] A systemic bank is indeed detected if the systemic shock realizes, but at that point its capital is fully exhausted and there is no further punishment that, under limited liability, can be imposed to the bankers investing in and managing it.

of using deposit insurance as an indirect (and arguably artificial) means of redistribution of wealth from the patient agents to the impatient agents.[18]

Finally, as already anticipated above, prudential regulation obliges banks to hold granular loan portfolios and establishes a minimum capital requirements of the form $e_{\xi t} \geq \gamma_t l_{\xi t}$. In the core sections of the paper we consider a time-invariant capital requirement, $\gamma_t = \gamma$. Time-varying (or state contingent) capital requirements are analyzed as an extension. In particular, we look at specific time-varying requirements when assessing the implications of moving to a regime of higher capital requirements in a gradual way. Macroprudential policies other than setting γ (e.g. altering the effective cost at which, given some value of r, banks can access non-equity funding) can also be analyzed as extensions.

4 Equilibrium analysis

This section derives the concepts and conditions relevant for the computation of equilibrium. In Subsection 4.1 we formalize bankers' portfolio decisions, which embed the choice on how much of their wealth to invest as bank capital, and the part of it going to the systemic or the non-systemic bank. This formalization is relevant for the writing of bankers' participation constraints in the contract problem that summarizes the way firms and banks establish the terms of their lending relationships, that we describe in Subsection 4.2. In Subsection 4.3, we discuss the solution to the contract problem, which embeds determining firms' production plans. The supply of bank capital resulting from the dynamics of bankers' wealth is described in Subsection 4.4. In Subsection 4.5 we provide the definition of equilibrium. Finally, our measure of social welfare W_t is introduced in Subsection 4.6 and further explained in the Appendix.

[18] The provision of deposit insurance at a zero flat rate premium is not essential for the results regarding banks' risk-shifting tendencies. Deposit insurance allows us to skip the determination of deposit rates and dealing with the issue of what happens with would-be bankers that invest their savings as deposits and lose them before becoming bankers. What is key to the results is the unobservability of banks' systemic orientation to everybody except banks' owners. The systemic bank is the great beneficiary of the deposit insurance subsidy, but absent deposit insurance, it will similarly benefit from pooling its deposits with those of the non-systemic bank, hence confronting deposit interest rates partly cross-subsidized by the other bank. Risk-shifting incentives would remain and the qualitative implications of the model would be very similar.

4.1 Bankers' portfolio problem

In our economy bankers' wealth, which is the only source of bank capital, earns endogenously determined scarcity rents. Yet, we model the market for bank capital as a perfectly competitive market in which bankers and banks take the relevant prices as given. Bankers who invest wealth as bank capital at t obtain some stochastic or state-contingent returns at $t+1$ that continuing bankers have the opportunity to reinvest as bank capital for at least one more period. Let v_{t+1} denote the (stochastic) marginal value of one unit of old bankers' wealth at the time of receiving the returns at $t+1$ (right before the banker learns whether he will remain active). Then active bankers' valuation at date t of a security j that pays a (stochastic) return R_{jt+1} at $t+1$ can be expressed as $\beta E(v_{t+1}R_{jt+1})$, where βv_{t+1} plays the role of a stochastic discount factor.

To write down the Bellman equation that determines v_t, we need to consider the wealth allocation opportunities available to an old banker at that date t. If the old banker converts into a worker, which happens with probability ψ, then his only alternatives are either to save the wealth in the form of bank deposits (earning a gross return $1+r$ at $t+1$) or to consume (in which case one unit of his wealth is worth just 1 at t). We assume that the agent's impatience (i.e. $\beta < 1/(1+r)$) together with the small probability of being a banker (or an entrepreneur) in the future imply that his optimal decision is to consume.

For the old banker who remains active at t (as well as for other impatient agents who become bankers at that date) the set of alternatives additionally includes investing in equity of the non-systemic bank (which yields a gross return R_{0t+1} per unit of wealth) and/or in equity of the systemic bank (which yields R_{1t+1}). Given the linearity in total wealth of this allocation problem (and our assumption about the best alternative for the non-continuing bankers), the Bellman equation for v_t can be expressed as:

$$v_t = \psi + (1-\psi)\max\{1, \beta \max\{(1+r)E_t(v_{t+1}), E_t(v_{t+1}R_{0t+1}), E_t(v_{t+1}R_{1t+1})\}. \tag{5}$$

This equation yields a number of properties for v_t and the various possible equilibrium allocations of bankers wealth. Specifically, continuing bankers' possibility of consuming their wealth at t implies $v_t \geq 1$ (although they will obviously not consume any wealth if

$v_t > 1$). Continuing bankers may decide to keep part of their wealth aside as bank deposits if $(1+r)E_t(v_{t+1}) \geq 1$ (so that they do not strictly prefer to consume their wealth) but R_{0t+1} and R_{1t+1} are small enough to have $(1+r)E_t(v_{t+1}) \geq \max\{E_t(v_{t+1}R_{0t+1}), E_t(v_{t+1}R_{1t+1})\}$. In fact, in equilibrium the last inequality will never hold strictly because in that case no banker would invest in bank capital and banks would not be able to give loans, which is a situation incompatible with equilibrium.[19]

For brevity, we are going to focus the presentation of equilibrium conditions in the main text on the case in which $\beta Q_t \equiv \beta \max\{E_t(v_{t+1}R_{0t+1}), E_t(v_{t+1}R_{1t+1})\} > \max\{1, \beta(1+r)E_t(v_{t+1})\}$, where Q_t is bankers' *required value-weighted return* on wealth. In this case, active bankers will devote their entire wealth to investing in equity of the non-systemic bank if $E_t(v_{t+1}R_{0t+1}) > E_t(v_{t+1}R_{1t+1})$, equity of the systemic bank if $E_t(v_{t+1}R_{1t+1}) > E_t(v_{t+1}R_{0t+1})$, or the two (or any of them) if $E_t(v_{t+1}R_{0t+1}) = E_t(v_{t+1}R_{1t+1})$.

To avoid problems interpreting the pooling equilibrium in which the systemic bank is intended to mimic the non-systemic bank (since otherwise we assume that it would be singled out and dissolved by the supervisor), we assume (and check in all the numerical results below) that the equity of the systemic bank is sufficiently attractive to bankers in equilibrium. In this case, the fraction of total bank capital that bankers allocate to the systemic bank satisfies $x_t \in [0, 1)$ for all t and we have $Q_t = E_t(v_{t+1}R_{0t+1})$ for all t.[20]

4.2 Firm-bank lending contracts

At the demand side of the market for bank capital, banks know that an arbitrary stochastic payoff P_{t+1} on one unit of bankers' wealth is acceptable to the bankers if and only if it pays to the bankers at least their required value-weighted return on wealth, $E(v_{t+1}P_{t+1}) \geq Q_t$. The contract signed between the non-systemic bank and each its funded entrepreneurs will

[19] In our economy loans are necessary for firms to produce and the properties of the production technology, together with the Walrasian determination of equilibrium wages, would tend to make supplying loans infinitely profitable if the amount of supplied loans tends to zero.

[20] It is possible to analytically show that having a small measure of bankers ($\phi \to 0$) or low risk-shifting incentives ($p_1 \to (p_0 - \varepsilon)/(1-\varepsilon)$) is sufficient to rule out equilibria with $x_t = 1$.

set $(\xi_{it}, k_{it}, n_{it}, l_{it}, B_{it}) = (0, k_t, n_t, l_t, B_t)$ where k_t, n_t, l_t, and B_t solve the following problem:

$$\begin{array}{cl} \max\limits_{(k_t,n_t,l_t,B_t,d_t,e_t)} & (1-p_0)[AF(k_t,n_t)+(1-\delta)k_t - B_t] \\ \text{s.t.} & E\{v_{t+1}[(1-p_0)B_t + p_0(1-\lambda)k_t - (1+r)d_t]\} \geq Q_t e_t, \\ & l_t = k_t + w_t n_t,\ \ l_t = d_t + e_t,\ \ e_t \geq \gamma l_t. \end{array} \tag{6}$$

This problem maximizes, subject to the constraints faced by the bank, the expected payoff of any of the funded entrepreneurs at the end of period t. The objective function reflects that the entrepreneur obtains the difference between the gross output $AF(k_t, n_t) + (1-\delta)k_t$ and the repayment B_t promised to the bank when his firm does not fail, and a zero payoff when the firm fails.

The first constraint in (6) is the bank's participation constraint. This constraint is written taking into account that the bank will hold a continuum of loans such as the one whose terms are optimized. Since non-systemic loans default in an uncorrelated manner, bankers' payoffs at $t+1$ are the gross repayments from the performing loans, $(1-p_0)B_t$, plus the income from the recovery of the depreciated physical capital of the non-performing loans (failed firms), $p_0(1-\lambda)k_t$, minus the bank's gross repayments to depositors, $(1+r)d_t$. The constraint says that bankers' expected value-weighted payoffs at $t+1$ must guarantee them the required value-weighted return Q_t per unit of initially contributed equity. The bank takes v_{t+1} and Q_t as *given* when solving (6).

The last three constraints in (6) reflect that (i) the loan is used to pay capital and labor in advance, (ii) the bank balance sheet imposes $l_t = d_t + e_t$, and (iii) the regulatory capital requirement imposes $e_t \geq \gamma l_t$.

Conveniently, the net payoffs $(1-p_0)B_t + p_0(1-\lambda)k_t - (1+r)d_t$ that bankers receive from the loans made by the non-systemic bank are, conditional on the information available at t, deterministic since they do not depend on the realization of the systemic shock at $t+1$ (and firms' idiosyncratic failure risk is perfectly diversified away). Hence this term can be taken out of the expectations operator in the bankers' participation constraint and the market-determined return on bank capital at non-systemic banks R_{0t+1} can also be treated as deterministic. From here, Q_t can be written as $E_t(v_{t+1})R_{0t+1}$, and $E_t(v_{t+1})$ can be eliminated from both sides of the participation constraint, which simplifies to:

$$(1-p_0)B_t + p_0(1-\lambda)k_t - (1+r)d_t \geq R_{0t+1}e_t. \tag{7}$$

This expression makes clear that R_{0t+1} and the wage rate w_t constitute the sole channels through which the firm-bank decisions $(k_t, n_t, l_t, B_t, d_t, e_t)$ are affected by the state of the economy at date t.[21]

If the systemic bank also operates in equilibrium (i.e. $x_t > 0$), it will simply mimic (except in setting $\xi_{it} = 1$ for all its funded firms), the terms of the lending contract and the capital structure decisions of the non-systemic bank.[22]

4.3 Production decisions resulting from lending contracts

It should the obvious that both (7) and the constraint associated with the minimum capital requirement that appear in (6) will be binding. So eventually the optimization problem involves six variables and four binding constraints. Moreover, the constraints can be conveniently used to reduce the problem to one of unconstrained optimization with just two variables: k_t and n_t. All the other variables can be found recursively using the binding constraints. So $l_t = k_t + w_t n_t$, $d_t = (1-\gamma)(k_t + w_t n_t)$, and $e_t = \gamma(k_t + w_t n_t)$. (7) and some further substitutions yield the following expression for the loan repayment:

$$B_t = \frac{1}{1-p_0}\{[(1-\gamma)(1+r) + \gamma R_{0t+1}](k_t + w_t n_t) - p_0(1-\lambda)k_t\}. \tag{8}$$

Intuitively, the loan repayment at $t+1$ must compensate in expected terms for the weighted average cost $(1-\gamma)(1+r) + \gamma R_{0t+1}$ of the funds $k_t + w_t n_t$ that the firm borrows from the bank at date t. The term $p_0(1-\lambda)k_t$ credits for the depreciated capital that the bank can recover if the firm fails.[23]

[21] If R_{0t+1} were not deterministic (say, because non-systemic firms also had some exposure to systemic shocks), the model would still be solvable but at a larger computational cost. The current property makes the system of equations that characterize equilibrium recursive by blocks.

[22] It can be shown that entrepreneurs are eventually indifferent between developing production in the systemic or non-systemic mode since in both cases the equilibrium lending contract implies that their net payoff is zero in all states. This is, as in neoclassical production theory, a result of perfect competition under constant returns to scale.

[23] The expression in curly brackets is divided by $1-p_0$ because B_t is only effectively paid when the firm does not fail.

Now, using (8) to substitute for B_t in the objective function of (6) gives rise to

$$\max_{(k_t,n_t)} (1-p_0)[AF(k_t,n_t)+(1-\delta)k_t]+p_0(1-\lambda)k_t-[(1-\gamma)(1+r)+\gamma R_{0t}](k_t+w_t n_t),$$

which is the reduced unconstrained maximization problem. This problem only differs from the typical profit maximization problem of perfectly competitive firms in static production theory in that (i) the production process is intertemporal and subject to failure risk, (ii) the expected gross output of the firm at $t+1$ is partly net output and partly depreciated capital, and (iii) the factors of production k_t and n_t are pre-paid at t using bank loans, so their cost (in $t+1$ terms) is scaled up by the weighted average cost of funds of the non-systemic bank, $(1-\gamma)(1+r)+\gamma R_0$.

Given the constant returns to scale of the technologies in the background, the objective function of the unconstrained optimization problem is homogeneous of degree one in (k,n). So similarly to all neoclassical models with this feature, obtaining finite non-zero solutions requires the value of the objective function to be zero at the optimum—so that if (k,n) is a solution, any proportional (k',n') is also a solution. Taking all variables except w_t and k_t as exogenous, the FOCs of the unconstrained problem (6) when evaluated at $n_t=1$ (which is the aggregate inelastic supply of labor) uniquely determine, for each value of R_{0t+1}, an equilibrium wage rate w_t and a physical capital to labor ratio k_t consistent with firm-bank optimization and labor-market clearing.[24] Specifically, we obtain

$$(1-p_0)[AF_k(k_t,1)+(1-\delta)]+p_0(1-\lambda)=[(1-\gamma)(1+r)+\gamma R_{0t+1}], \tag{9}$$

which determines a k_t for each R_{0t+1}, and

$$(1-p_0)AF_n(k_t,1)=[(1-\gamma)(1+r)+\gamma R_{0t+1}]w_t, \tag{10}$$

which recursively determines a w_t for each R_{0t+1} and k_t.

The above expressions are very transparent on one way in which bank frictions can affect the real sector in this economy: by increasing the competitive pricing of the bank loans that

[24] Notice that the possible presence of firms operated in the systemic mode does not alter the aggregation implicit in this argument since they mimic the (k,n) decisions associated with the lending contract of the non-systemic bank.

firms use to finance their factors of production. For given capital requirement γ, increasing the required rate of return on bank capital R_{0t+1}, increases such a price, pushing firms to reduce their scale, which, after the adjustment in the equilibrium wage rate, implies that both k_t and w_t fall. The same effects follow from an increase in γ, for given R_{0t+1}. But of course these are pure partial equilibrium effects since the main feature of this model is the endogeneity of R_{0t+1}.

The demand for bank capital that emanates from the previous discussion is

$$e_t = \gamma(k_t + w_t), \tag{11}$$

which is decreasing in R_{0t+1}. Hence determining the equilibrium path for R_{0t+1} is conceptually as simple as looking at the supply side of the market for bank capital and making sure that such market clears at all dates. The supply of bank capital is the result of the process of wealth accumulation by bankers (and the portfolio decisions already discussed in Subsection 4.1).

4.4 The supply of bank capital

For the purposes of this subsection, think of e_{t+1} as the aggregate *supply* of bank capital at date $t+1$—in equilibrium supply and demand are equal and we save on notation by using the same variable to represent both of them. Along a full reinvestment path, e_{t+1} coincides with the total wealth of active bankers at the beginning of period $t+1$, and such wealth is made up of two components: (i) the capitalized value $\phi(1+r)w_t$ of the labor income earned by the currently active bankers in the prior date (and that all of them are assumed to find it optimal to save), and (ii) the gross returns on the wealth $(1-\psi)e_t$ that continuing bankers invested as bank capital at date t.[25] This results in the following law of motion for e_{t+1}:

$$e_{t+1} = \phi(1+r)w_t + (1-\psi)[(1-x_t)R_{0t+1} + x_t R_{1t+1}]e_t, \tag{12}$$

[25] See Appendix A for the statement of all equilibrium conditions for the general case in which bankers active at any given date may find it optimal to consume part of their wealth or to keep part of it inverted as bank deposits. For simplicity, the assumption that bankers and would-be bankers save all their labor income will not be relaxed, but we will check that under the parameterizations that we explore such behavior is always optimal.

where, as previously defined, $x_t \in [0, 1]$ is the fraction of total bank capital allocated to the systemic bank at date t.

As already commented, the realized gross return on equity at the non-systemic bank R_{0t+1} is independent of the realization of the systemic shock at $t+1$, while the realized gross return on equity at the systemic bank R_{1t+1} is not. Under most reasonable parameterizations, if the systemic shock does not realize ($\varepsilon_{t+1} = 0$) one unit of capital of the systemic bank yields some $R^{1-\varepsilon}_{1t+1} > R_{0t+1}$, while if it realizes ($\varepsilon_{t+1} = 1$) the bank becomes insolvent and, by limited liability, bankers' gross return on equity becomes $R^{\varepsilon}_{1t+1} = 0 < R_{0t+1}$. In particular, using the fact that the systemic bank mimics the non-systemic bank in every decision except its unobservable systemic orientation and taking into account the differences in default rates experienced by the non-systemic bank, p_0, and the systemic bank, p_1, when $\varepsilon_{t+1} = 0$, it is possible to find the following expression for $R^{1-\varepsilon}_{1t+1}$ in terms of R_{0t+1} and other parameters and variables of the model:[26]

$$R^{1-\varepsilon}_{1t+1} = \frac{1-p_1}{1-p_0} R_{0t+1} + \frac{1}{\gamma}\frac{p_0 - p_1}{1-p_0}[(1-\gamma)(1+r) - (1-\lambda)\frac{k_t}{k_t + w_t}]. \tag{13}$$

From the point of view of date t, R_{1t+1} is a dichotomous random variable that solely depends on the realization of ε_{t+1}. This implies that, from the same perspective, e_{t+1} is also a dichotomous random variable, and the aggregate uncertainty embedded in (12) can be better visualized by re-expressing it as:

$$e_{t+1} = \begin{cases} e^{1-\varepsilon}_{t+1} = \phi(1+r)w_t + (1-\psi)[(1-x_t)R_{0t+1} + x_t R^{1-\varepsilon}_{1t+1}]e_t, & \text{if } \varepsilon_{t+1} = 0, \\ e^{\varepsilon}_{t+1} = \phi(1+r)w_t + (1-\psi)(1-x_t)R_{0t+1}e_t, & \text{if } \varepsilon_{t+1} = 1. \end{cases} \tag{14}$$

From here, it is possible to reconsider (5) in order to express mathematically the compatibility of particular values of x_t with bankers' optimal portfolio decisions. Specifically, recalling that we want to abstract from equilibria with $x_t = 1$, we can establish that supporting equilibria with $x_t \in [0, 1)$ only requires:

$$[(1-\varepsilon)v(e^{1-\varepsilon}_{t+1}) + \varepsilon v(e^{\varepsilon}_{t+1})]R_{0t+1} - (1-\varepsilon)v(e^{1-\varepsilon}_{t+1})R^{1-\varepsilon}_{1t+1} \geq 0. \tag{15}$$

[26] The expression makes clear that, when the capital requirement γ is lower than the depreciation of physical capital in failed projects λ, the assumption that $p_0 > p_1$ implies $R^{1-\varepsilon}_{1t+1} > R_{0t+1}$. In our parameterizations below, the socially optimal capital requirement γ^* is always well below λ.

In fact if this inequality holds with equality, then any $x_t \in [0,1)$ is compatible with bankers' optimization, but if the inequality is strict, then only the corner solution $x_t = 0$ is consistent with bankers' optimization. To capture this formally, the following complementary slackness condition must be imposed:

$$\{[(1-\varepsilon)v(e_{t+1}^{1-\varepsilon}) + \varepsilon v(e_{t+1}^{\varepsilon})]R_{0t+1} - (1-\varepsilon)v(e_{t+1}^{1-\varepsilon})R_{1t+1}^{1-\varepsilon}\}x_t = 0. \tag{16}$$

4.5 Definition of equilibrium

A *full-reinvestment equilibrium* is (i) a stationary law of motion for the state variable e on a bounded support $[\underline{e}, \overline{e}]$ and (ii) a tuple $(v(e), x(e); k(e), w(e),\ R_0(e), R_1^{1-\varepsilon}(e))$ describing the key endogenous variables as functions of $e \in [\underline{e}, \overline{e}]$, such that all the sequences $\{e_t\}_{t=0,1,...}$ and $\{v_t, x_t, k_t, w_t, R_{0t+1}, R_{1t+1}^{1-\varepsilon}\}_{t=0,1,...}$ generated by each of them are compatible at all dates with:

1. Optimization by price-taking workers, bankers, entrepreneurs, banks, and firms.

2. The clearing of all relevant markets.

3. The investment as bank capital of all the wealth available to the active bankers.

If the economy supports a full-reinvestment equilibrium, then its state at any date t can be summarized by just one variable: the total wealth available to the active bankers e_t, which they will find it optimal to invest as bank capital. The law of motion of such variable is given by (14), which makes clear that the stochastic evolution of e_t is driven by the realization or not of the systemic shock at the end of every period. The tuple $(v(e), x(e); k(e), w(e),\ R_0(e), R_1^{1-\varepsilon}(e))$ mentioned in the definition above contains functions of e that can be used to determine all other equilibrium variables for each given value of the state variable $e_t = e$ in the relevant support $[\underline{e}, \overline{e}]$. In particular, the marginal value of bank capital can be found as $v_t = v(e_t) \geq 1$, the fraction of bank capital allocated to the systemic bank as $x_t = x(e_t) \in [0,1)$, the amount of physical capital used by firms as $k_t = k(e_t) \geq 0$, the wage rate as $w_t = w(e_t) \geq 0$, the (anticipated) return of bank capital invested in non-systemic loans as $R_{0t+1} = R_0(e_t) \geq 1+r$, and the (anticipated) return of bank capital invested in systemic loans if $\varepsilon_{t+1} = 0$ as $R_{1t+1}^{1-\varepsilon} = R_1^{1-\varepsilon}(e_t) \geq 0$, All the conditions that the functions in the tuple $(v(e), x(e); k(e), w(e), R_0(e), R_1^{1-\varepsilon}(e))$ must satisfy at each value of e in order to be compatible with requirements 1-3 in the definition have been already expressed as equalities or inequalities in the equations of previous sections. Appendix A extends these conditions to deal with equilibria in which active bankers may find it optimal to consume part of their available wealth or to invest it in bank deposits (rather than as bank capital) at some dates.

Appendix A also describes the numerical solution method that we use to compute the equilibrium associated with the parameterizations explored in the quantitative part of our analysis. In essence, we use non-linear recursive techniques based on a discretization of the state space, $[\underline{e}, \overline{e}]$, iterating the value function $v(e)$ that describes the marginal value of one unit of bankers' wealth when the total wealth available to bankers is e (equation (5)), and using the system made up of all the other relevant equations to find consistent equilibrium values of the other elements of $(v(e), x(e); k(e), w(e), R_0(e), R_1^{1-\varepsilon}(e))$.

Importantly, since low values of e imply a lower supply of bank capital, the function $v(e)$ will be decreasing in e, reflecting that in this economy one unit of bankers' wealth is more valuable when bankers' wealth is overall more scarce. Indeed the decline of $v(e)$ with e is strict, insofar as bankers keep devoting their wealth to banks' equity. In this case, a higher e implies that banks have more capital and can give more loans, which in turn implies that firms can employ more physical capital (and/or more labor). But, like in the neoclassical growth model, the equilibrium returns to investment are marginally decreasing (since the supply of labor is fixed), making the value of bank lending and, hence, the rents appropriated by the bankers (as exclusive suppliers of bank capital) also marginally decreasing in e.[27]

To understand the intuition driving the equilibrium systemic risk-taking decisions described by $x(e)$, the crucial equations are the counterparts of (15) and (16). As previously mentioned (13) implies $R_{1t+1}^{1-\varepsilon} > R_{0t+1}$, so satisfying (15) requires a sufficiently large value of the term $\varepsilon v(e_{t+1}^{\varepsilon})$, which intuitively captures the expected marginal value of protecting one unit of bankers' wealth from the systemic shock (i.e. investing it as capital of the non-systemic bank). Indeed, the capital invested in the systemic bank gets completely lost when

[27] This result also arises, with identical intuition, in e.g. Gertler and Kiyotaki (2010).

the systemic shock realizes, while the bank capital invested in the non-systemic bank is more than just preserved (since it gets capitalized at the deterministic gross return on equity R_{0t+1}). But the marginal value of bankers' wealth if the systemic shock realizes, $v(e^{\varepsilon}_{t+1})$, is endogenously affected by the fraction of bank capital invested in a systemic manner x_t. By (14), a larger x_t implies, other things equal, a larger aggregate loss of bank capital when the shock occurs, and hence a lower e^{ε}_{t+1} and a larger $v(e^{\varepsilon}_{t+1})$. This establishes a self-equilibrating mechanism for the determination of x_t. (16) embeds the no-arbitrage relationship (or indifference condition) required for producing an interior $x_t \in (0,1)$, which is possible thanks to the just-commented equilibrating force.

4.6 Social welfare

The natural measure of social welfare W_t in this economy is the expected present value of the aggregate net consumption flows that it provides to the various final agents.[28] This measure can be obtained and decomposed in various forms, depending on the dimension along which the relevant overall aggregation is performed. One can infer the net consumption flow that corresponds to each class of agents in each date t and aggregate across agents. Alternatively, one can just look at the differences between the aggregate quantity of the consumption good the economy produces at the end of a typical period and the quantity which is "reutilized" as a factor of production (physical capital) for next-period production. Appendix B provides an explicit expression for W_t and an associated flow measure of welfare ω_t, which is explained using two possible decompositions.

Following convention, we will describe the welfare gains and losses below in reference to percentual differences in the *certainty-equivalent permanent aggregate net consumption flow* that would produce such welfare, which can be calculated as simply $(1-\beta)W_t$.

[28] See Appendix B for a justification of the use of impatient agents' discount factor β in the discounting of the relevant consumption flows.

5 Numerical results

5.1 Calibration

Table 1 contains the parameters chosen for a first, tentative calibration of the model. We generate most of the results presented below under these baseline parameters. We have tried numerous other parameter configurations to check the performance of the solution method and the generality of the qualitative results, in both cases with positive results.

Table 1
Baseline parameter values
(One period is one year; all rates are yearly rates)

Discount rate of the patient agents	r	0.02
Discount factor of the impatient agents	β	0.96
Total factor productivity	A	2
Physical capital elasticity	α	0.3
Depreciation rate in successful firms	δ	0.05
Depreciation rate in failed firms	λ	0.35
Idiosyncratic default rate of non-systemic firms	p_0	0.03
Idiosyncratic default rate of systemic firms	p_1	0.018
Probability of a systemic shock	ε	0.03
Bankers' exit rate	ψ	0.20
Fraction of wages devoted to forming new bank capital	ϕ	0.05

The model is very parsimonious: it has the 11 parameters listed in Table 1 (plus the capital requirement γ, when taken as given rather than set so as to maximize social welfare) and a single binary iid aggregate shock—the systemic shock—whose probability of occurring is one of the 11 parameters. The discount rate of the patient agents r is chosen equal to 2% to capture a situation with low real interest rates such as the one that characterized most developed economies in the years leading to the current crisis. The discount factor of the impatient agents embeds a discount rate which is approximately twice as large as r. In the literature on financially-constraint risk-neutral agents it is standard to assign values of this order of magnitude or even larger to their discount rate. The value for the total factor productivity parameter A is inconsequential, except for setting the scale of all the variables in levels. $A = 2$ produces levels in a range from 0 to 100 which are just easy to report in the tables below.

The elasticity of physical capital in the production function α is fixed according to standard macro practice, so as to produce a share of labor income in GDP of about 70%. The depreciation rates of physical capital in successful and failing firms, δ and λ, are chosen so as to match an aggregate physical capital to GDP ratio in the range of 3 to 4 as well as a loss-given-default (LGD) for bank loans of about 45%, which is the LGD fixed for unrated corporate exposures in the standardized approach of Basel II.[29]

The values of p_0, p_1, and ε are set so as to have sufficient potential room of risk shifting and for significant aggregate losses due to it. The current choices are compatible with the conditions $p_1 < p_0 < (1-\varepsilon)p_1+\varepsilon$ established in section 3.2. They imply *expected* default rates in the range from 3% (if all firms were non-systemic) to 4.7% (if all firms were systemic). The probability of a systemic shock (situation in which all systemic firms fail simultaneously) is set at 3%, so that this shock occurs on average once in every 33 years. Arguably all these choices are above what recent historical experience in Western economies would suggest as realistic. They are made at this point to help extract a first set of *visible* qualitative lessons from our analysis. All in all the results below are sufficiently important to suggest that the results would also be quantitatively significant under a more conservative choice of these three parameters. We plan to check this in the immediate future.

Finally, the current choices for the values of the bank capital dynamics parameters, ψ and ϕ, is also very tentative. The rate ψ at which bankers convert into workers and stop providing their accumulated wealth as bank capital implies that they have an average active life of 5 years and, according to the results reported below, a flow of bank capital exits equivalent to payouts of between 5% and 15% of the stock of bank capital per year. The value of ϕ implies that bankers' represent 5% of the efficiency units of labor in the economy,

[29] To explain why $\lambda = 0.35$ produces a LGD of 45%, notice that loans in this model also finance firms' wages (what empirically might correspond to funding their working capital) and the recoveries obtained by banks on that part of the loans in zero when a firm fails.

so that their yearly labor income provides additions to the aggregate wealth available to active bankers equivalent to 5% of total labor income per period.[30]

5.2 Understanding the results

The numerical method used for solving the model is described in Appendix A. It is based on value function iteration for obtaining $v(e)$, and uses the relevant set of equilibrium conditions, without linearizing or approximating them in any other form, except for interpolation purposes. We frequently came across situations in which it was necessary to deal with the possibility of equilibria without full reinvestment of bankers' wealth, although those situations were typically confined to values of e out of the support of the corresponding ergodic distribution of this variable. It was also relatively frequent to find solutions with $x_t = 0$. So all the results reported below correspond to the extended set of equilibrium conditions, provided in Appendix A that accommodates these possibilities. No problem of multiplicity of equilibria was detected.

For illustration purposes, we will systematically compare the results obtained with a capital requirement of 7% (γ=0.07) with those obtained with the requirement that, under the baseline parameters in Table 1, happens to maximize social welfare W_t when evaluated at a period t in which the economy is at its pseudo-steady state. Such optimal minimal capital requirement is 14% (γ=0.14). Figure 1, generated by solving the model for a grid of values of γ, and computing the relevant W_t along sufficiently many generated histories of the economy anticipates this first result.

[30] In what we call the *pseudo-steady state* of the economy (the state reached after sufficiently many periods without experiencing the systemic shock), the increase in active bankers' available wealth due to ϕ exactly compensates the wealth taken out by the fraction ψ of non-continuing bankers. This makes the effective supply of bank capital in the pseudo-steady state to remain constant until a systemic shock occurs.

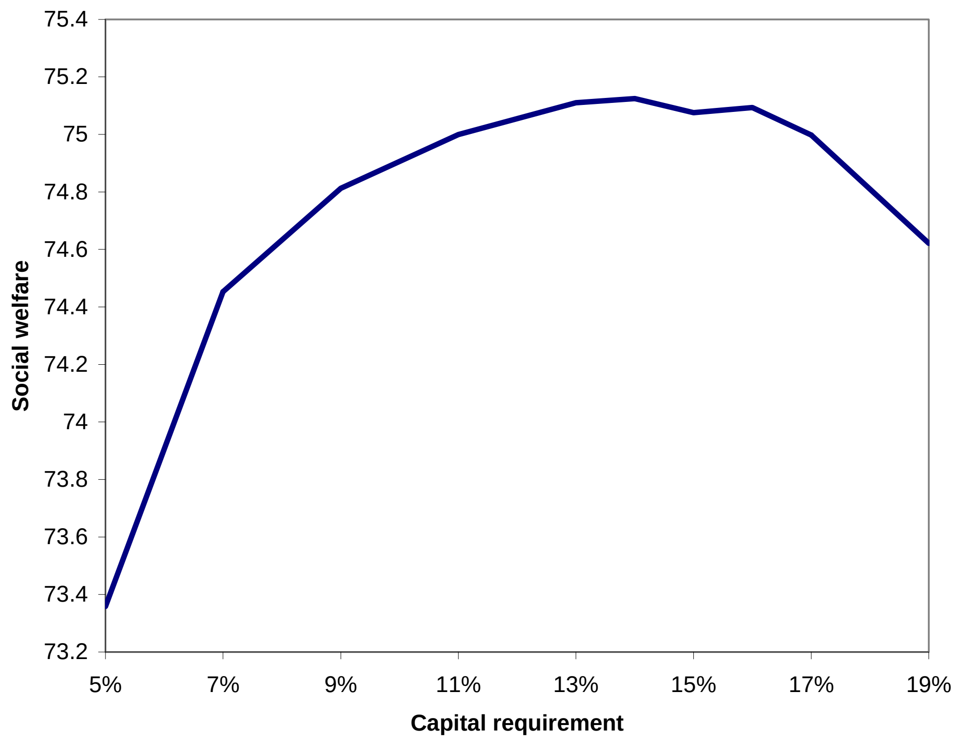

Figure 1 Social welfare W as a function of the capital requirement γ

Before trying to understand the result, it is worth looking at the nature of the solution to our model. Figure 2 depicts two central objects in our definition of equilibrium: the function $v(e)$ that describes the marginal value of one unit of bank capital, and the function $x(e)$ that describes the fraction of bank capital invested in the systemic bank (or, equivalently, devoted to make systemic loans). Both functions are depicted in Figure 2 for the union of the ranges of e which are relevant in the economies with $\gamma = 0.07$ (low capital requirement) and $\gamma = 0.14$ (optimal capital requirement)—in the economy with low γ total bank capital tends to fluctuate over a wider but lower range of values than in the economy with optimal γ.

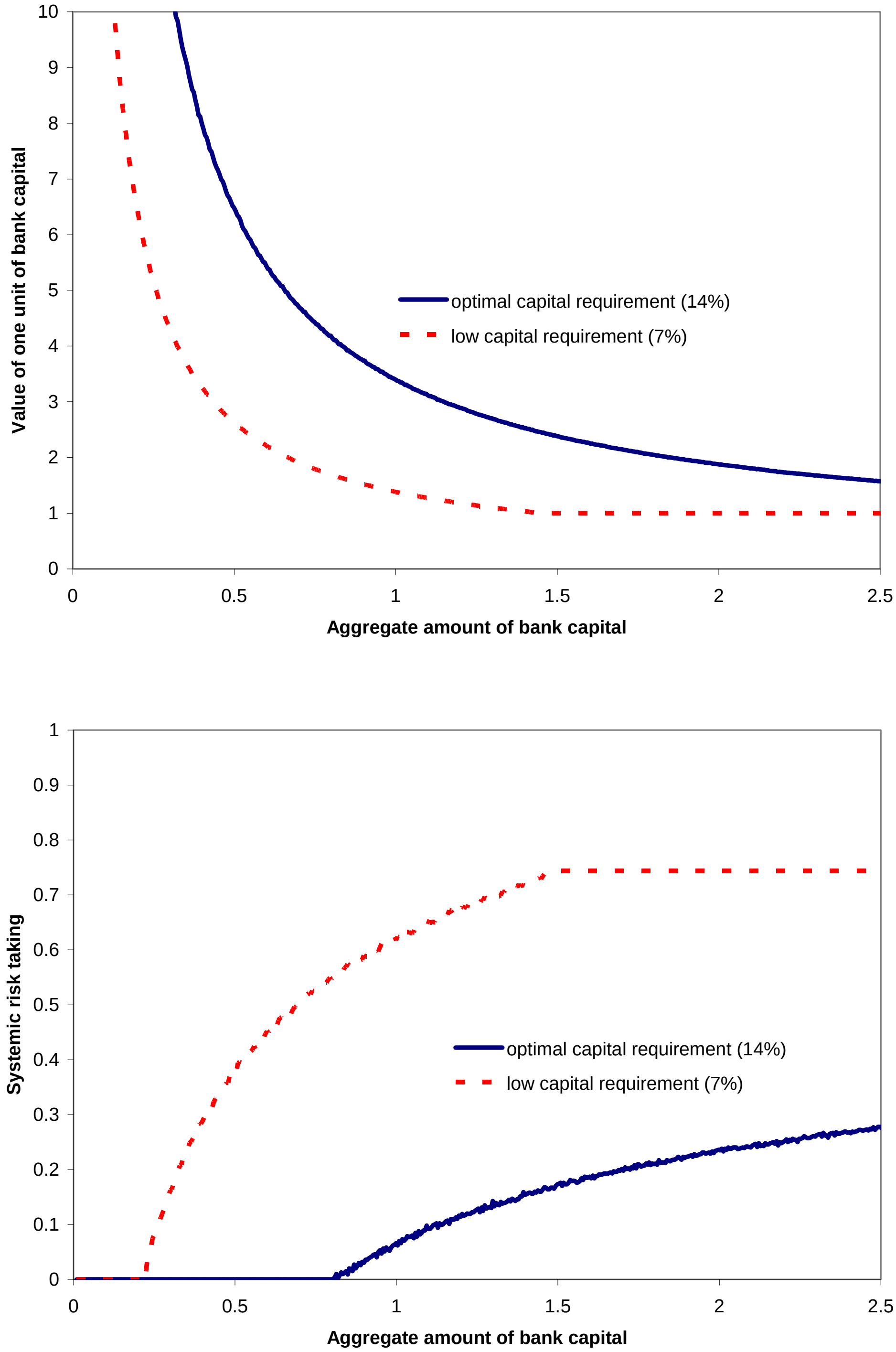

Figure 2 $v(e)$ and $x(e)$ under low and optimal capital requirements (CR)

The results tell, quite intuitively, that the greater scarcity of bank capital associated with the higher capital requirement implies that it gets a higher marginal value at every level of capital e. More importantly, systemic risk-taking, which is non-decreasing in e (and strictly increasing in e whenever its value is interior), is lower at every e when the capital requirement γ is 14% than when it is 7%. In fact the monotonicity of the effect of γ on the schedule $x(e)$ holds for all the range of capital requirements that we have tried. The central intuition for this result is that the impact of γ on $v(e)$ (i.e. on its position and, possibly more importantly, on its slope) affects the incentives of bankers to guarantee that (part of) their capital survives a systemic shock (in opposition to the alternative of appropriating possible risk-shifting gains but exposing themselves to the loss of all capital if the shock arrives).

The rationale for bankers systemic risk choices as reflected in $x(e)$ and its interaction with the (endogenous) dynamics of bank capital can be further explained by looking at Figures 3 and 4, each corresponding to one of the two capital requirements that we are comparing. The structure and qualitative results emanated from each figure are very similar. The solid line in the first panel of each of them shows the mapping from the amount of bank capital in one period, e_t, to its amount in the next period, e_{t+1}, if the systemic shock does not occur "at the end" of period t. This schedule is strictly increasing except when the aggregate amount of bank capital is large enough for continuing bankers to start to consume part of it as a voluntary dividend (an option considered in the extended definition of equilibrium described in Appendix A).[31]

The dashed downward sloping curve in the figures shows the mapping from e_t to e_{t+1} when the systemic shock occurs. The vertical distance between the solid and the dashed curves measures the bank capital lost when the economy is hit by the systemic shock. The loss is bigger not only in absolute but also in relative terms because the extent of systemic risk taking ($x(e)$ in Figure 2) is higher for higher values of e. Indeed, for sufficiently low values of e, we have $x(e) = 0$, which explains why the two key curves in the top panel of Figures 3 and 4 intersect.

[31] This explains the flat portion of the solid curve depicted in the first panel of Figure 3.

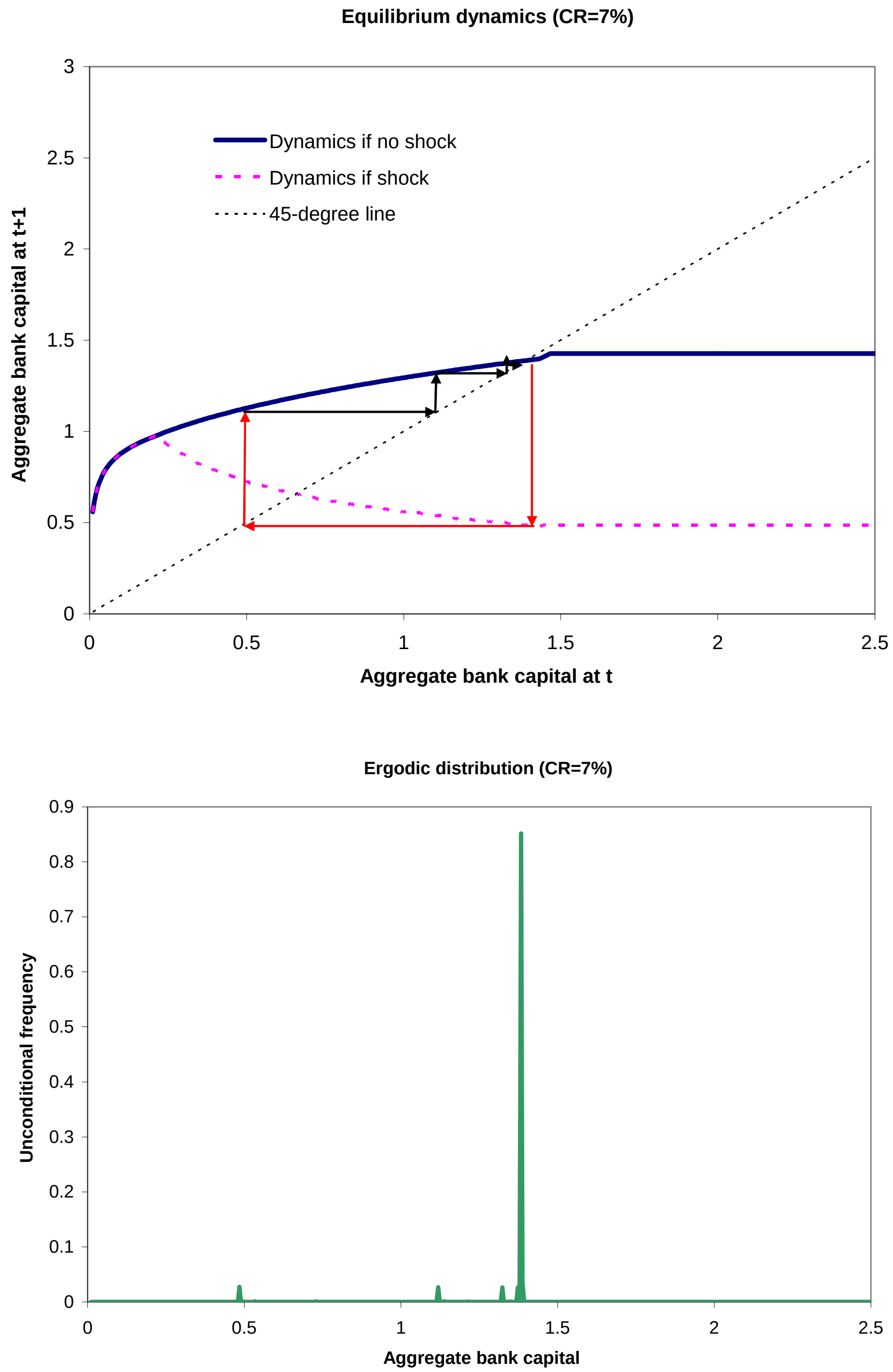

Figure 3 Equilibrium dynamics with low capital requirements (7%)

The points where the solid no-shock phase diagrams intersect the 45-degree line in each figure identify the pseudo-steady states. Interestingly, the value of e in the pseudo-steady state is the point in the ergodic support of e (values possibly reached with non-negligible probability along sufficiently long histories of the economy) associated with the highest level of systemic risk-taking . This is also the point in which the consequences of the realization of the systemic shock are more devastating, since it implies the largest loss of bank capital and is followed by the largest contraction of credit. The arrows on Figures 3 and 4 identify the path of crisis and recovery for the particular (but most frequent!) situation in which the economy fully returns to its pseudo-steady state without suffering a second systemic shock (recall that systemic shocks are iid and occur with a probability of just 3% per period in our calibrations). With both $\gamma = 0.07$ and $\gamma = 0.14$ our economy fully recovers in a minimum of 5 periods (the last period prior to recovery is very close to the pseudo-steady state and is hardly visible in the figures).

The second panels in Figure 3 and 4 depict the relative frequencies with which different values of e are visited along sufficiently long histories of the economy. This generates what we call the ergodic (unconditional) probability distribution of e. Consistently with the aggregate shock being so little frequent, our economy spends most of the time (about 80% or more of it) in the pseudo-steady state. However, when the systemic shock comes the consequences are very sizeable. Other points with positive frequencies correspond the various recovery (or secondary and subsequent crisis and recovery) paths that may occur.

For the comparison of the economies with $\gamma = 0.07$ and $\gamma = 0.14$ in welfare terms, the most relevant aspects reflected in previous figures are (i) the amount of systemic risk taking $x(e)$ over the range of relevant values of e and (ii) the amplitude of such range. It follows from the bottom panel of Figure 2 and the top panels of Figures 3 and 4 that both (i) and (ii) are much bigger with $\gamma = 0.07$. When disentangling the sources of the welfare gains associated with adopting the optimal capital requirement, we will refer to the efficiency gains from having lower $x(e)$ at the pseudo-steady state as *static gains* and the gains that accrue at any other state (either due to inefficient risk-shifting or the lower level of activity caused by the low values of e that follow a first systemic shock) as *dynamic gains*.

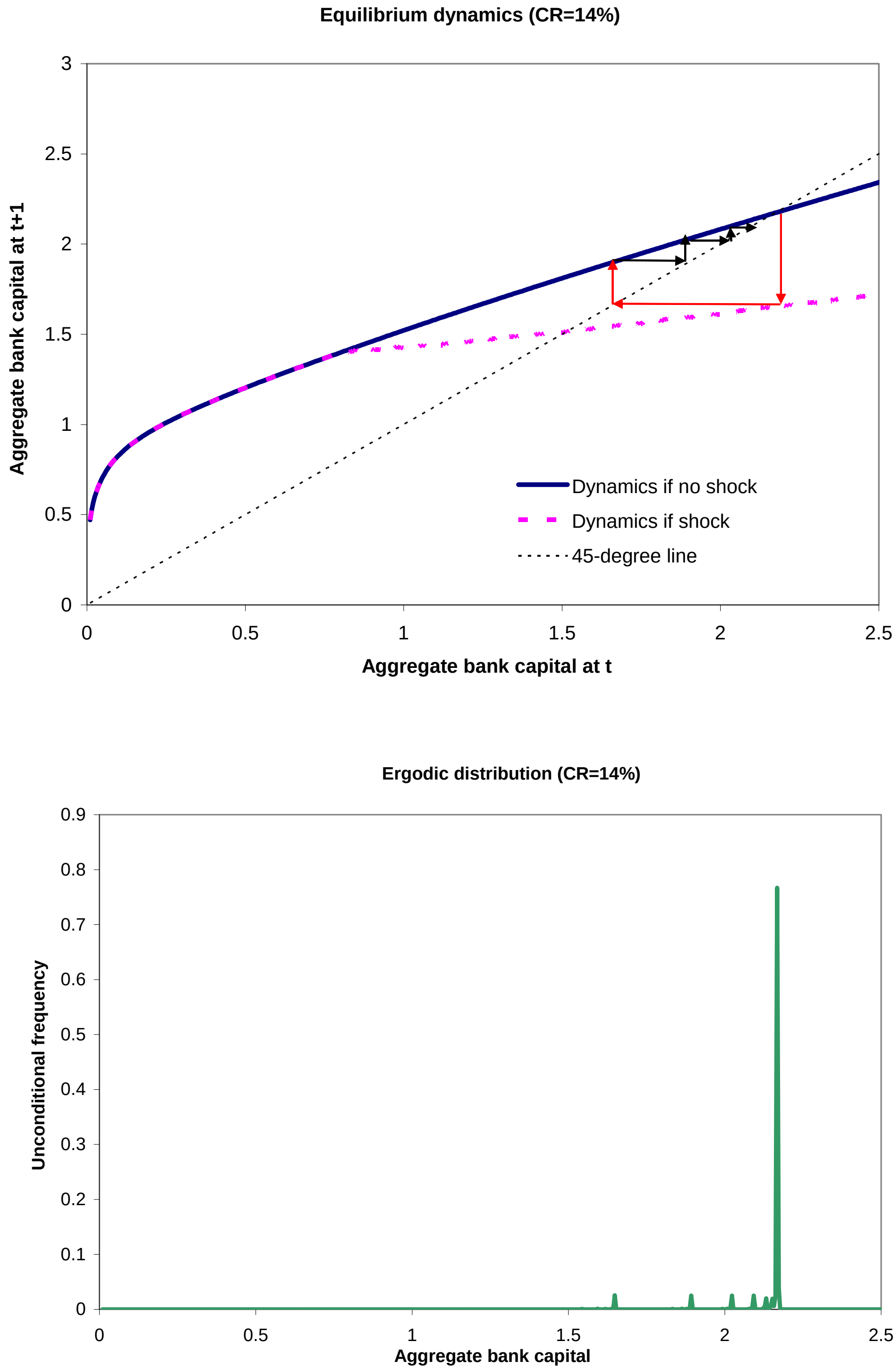

Figure 4 Equilibrium dynamics with optimal capital requirements (14%)

5.3 Quantitative details

We next report the quantitative details of the equilibria underlying the figures that we have just described. To to this, Table 2 reports a number of relevant statistics, including the overall welfare reached by each economy in its corresponding pseudo-steady state (PSS), W_{PSS}, which is the criterion that we have maximized to find the optimal capital requirement of 14%. The table reports welfare in terms of the certainty-equivalent permanent aggregate net consumption flow that would produce W_{PSS} (which can be computed as $(1-\beta)W_{PSS}$).

We find that moving from the low to the optimal capital requirement increases our overall measure of social welfare in an amount equivalent to a perpetual increase of 0.9% in aggregate net consumption. The main reason behind this increase in welfare is that the move to higher capital requirements reduces the fraction of systemic loans in the pseudo-steady state from roughly 72% to roughly 25%. In terms of the decomposition of the welfare gains that we suggested above, we find that 65% of the gain is due to "static gains"—direct gains from reducing inefficient risk-shifting in normal periods—and the remaining 35% is due to "dynamic gains," i.e. the reduction in the losses incurred during the periods it takes for the economy to return to its pseudo-steady state after a systemic shock.

Table 2 also reports several macroeconomic and banking aggregates, and ratios derived from them. Some columns report the values for the $\gamma = 0.07$ economy, other for the $\gamma = 0.14$ economy, other report the differences between the statistics of both economies, and finally there are columns (at the bottom of the table) in which we look at variations within each economy when, starting from the PSS, they are hit by a systemic shock. Specifically, we report PSS values, the values in the first period after the shock (if no second shock occurs) and the differences across these two states. The goal there is to show how such differences differ across the economies with low and optimal capital requirements.

Importantly, moving capital requirements from 7% to 14% brings about severe declines in the pseudo-steady state values of macroeconomic aggregates such as bank credit, GDP, and wages, which fall on average by 22%, 7%, and 9%, respectively. Our analysis evidences that, in the presence of systemic risk, the pseudo-steady state values of these variables are pretty bad proxies of social welfare since they do not reflect the very significant losses suffered when such risk materializes. Indeed, as reflected in the table, the fall in aggregate net consumption in the year that follows a systemic shock is (relative to the pseudo-steady state level) of 11.7% with capital requirements of 7% and of only 2.8% with capital requirements of 14%. Similarly, the falls in bank credit, GDP, and wages in the year after a shock are of 65%, 30%, and 37% with capital requirements of 7%, while it is of 24%, 9%, and 11% with capital requirements of 14%.

All in all the results suggest that the quantitative implications of capital requirements can be quite sizeable. They also point to the fact that setting capital requirements optimally must necessarily respond to an economic risk-management logic: caring about inefficient systemic risk-taking and its normally-invisible threat to macroeconomic stability. Checking the value of standard macroeconomic variables (such as GDP or wages) in the PSS, especially along the path in which the systemic shock does not realize, gives a very bad indication of the convenience of setting high capital requirements. Optimal capital requirements always seem to have "large costs" in terms of these variables.

The comparison between the economies with $\gamma = 0.07$ and $\gamma = 0.14$ in terms of banking and macro-financial ratios at the relevant PSS (such as the ratio of bank credit to GDP) is also very interesting. It tells mainly about the correlation between the underlying endogenous level of systemic risk-taking and several indicators of "financial exuberance" (e.g. bank credit to GDP). This correlation (very much in line with the perceptions of the early proponents of a macroprudential approach to bank regulation) is endogenously generated by the model.

Table 2. Provisional quantitative results

Welfare

	γ=7%	γ=14%	% Change
Certainty-equivalent aggregate net consumption (CEC)	2.978	3.005	0.9%
Expected aggregate net consumption in PSS (static)	2.987	3.008	0.7%
Aggregate net consumption in normal times if no shock	3.183	3.065	-3.7%
CEC off PSS (dynamic)	2.908	2.991	2.8%
Frequency of normal times	88.6%	82.9%	

Pseudo-steady state level statistics

	γ=7%	γ=14%	% Change
Expected aggregate net consumption	2.99	3.01	0.7%
Expected GDP	4.45	4.14	-7.1%
Aggregate net consumption if shock does not realize	3.18	3.07	-3.7%
GDP if shock does not realize (=Baseline GDP)	4.55	4.17	-8.4%
Bank credit	19.63	15.41	-21.5%
Physical capital	16.54	12.62	-23.7%
Wages	3.09	2.80	-9.4%
Bank capital	1.39	2.17	56.4%
Cost of deposit insurance if shock realizes	-5.66	-1.33	-76.4%

Basic real and financial ratios in PSS	γ=7%	γ=14%	Difference
Wages/Baseline GDP	0.679	0.671	-0.007
Physical capital/Baseline GDP	3.633	3.026	-0.607
Systemic risk-taking (fraction of systemic loans)	0.716	0.250	-0.466
Loan spread (loan rate - deposit rate)	0.017	0.035	0.018
Bank credit/Baseline GDP	4.312	3.697	-0.615
Bank capital/Baseline GDP	0.305	0.520	0.216
Return on equity at non-systemic banks	0.051	0.158	0.107
Bankers' payouts/Bank capital	0.114	0.066	-0.048
Marginal value of bank capital	1.046	1.760	0.713

Year after-the-shock statistics (when shock arrives in normal times)

		γ=7%			γ=14%	
	PSS	Year after shock	% Change	PSS	Year after shock	% Change
Expected aggregate net consumption	2.99	2.64	-12%	3.01	2.92	-3%
Expected GDP	4.45	3.13	-30%	4.14	3.77	-9%
Bank credit	19.63	6.92	-65%	15.41	11.77	-24%
Physical capital	16.54	4.98	-70%	12.62	9.28	-26%
Wages	3.09	1.94	-37%	2.80	2.49	-11%
Minimal number of years to full recovery		5			5	

6 Work ahead

We need to fine tune the details of our calibration and complete the discussion of the results. Additionally there are a number of other issues that we plan to discuss in full extent in the complete version of the paper:

- Transitional dynamics from moving γ and impact on the welfare assessment of an increase in capital requirements.
- Value (and limits to the value) of gradualism when introducing higher capital requirements.
- Assessment of capital requirements intended to have a counter-cyclical effect.
- Complementarity between capital requirements and "credit policies" (e.g. policies directed to make credit relatively more expensive in normal times and/or relatively less expensive in crisis times).
- Other macroprudential policies.

7 Concluding remarks

The paper describes a dynamic general equilibrium model of endogenous systemic risk-taking. The model incorporates a meaningful definition of systemic risk and allows a formal assessment of macroprudential policies in general and bank capital requirements in particular. Importantly, the model embeds an internally consistent welfare metrics according to which normative analysis can be performed. The preliminary results suggest that capital requirements have qualitatively and quantitatively important effects on systemic risk-taking, standard macroeconomic and banking indicators, and welfare. Although capital requirements appear to have a significant cost in terms of standard macroeconomic variables measured in "normal times," the results suggest that there is a unique socially optimal level for the capital requirements, that finding it out requires a carefully calibrated formal model, and the social welfare at stake can be quite large.

Appendix

A Solution method

The numerical solution procedure used in order to compute the equilibrium of the model can be described as follows:

1. Create a grid $\{e_i\}$, with $i = 1, 2, ...N$ (with some large N), over a range of values that includes the conjectured relevant range $[\underline{e}, \overline{e}]$ of the state variable

2. For each point in the grid, define $(k(e_i), w(e_i), R_0(e_i))$ as the (unique) non-negative (k_i, w_i, R_{0i}) that, for the given e_i, solve the following version of equilibrium conditions (9)-(11):

$$(1 - p_0)[AF_k(k_i, 1) + (1 - \delta)] + p_0(1 - \lambda) - [(1 - \gamma)(1 + r) + \gamma R_{0i}] = 0,$$

$$(1 - p_0)AF_n(k_i, 1) - w_i[(1 - \gamma)(1 + r) + \gamma R_{0i}] = 0,$$

$$\gamma(k_i + w_i) - e_i = 0.$$

3. Identify, if it exists, the point in the grid j for which $R_{0j} \geq 1 + r$ but $R_{0j+1} < 1 + r$.

   (a) For $i < j + 1$ set $\widehat{e}_i = e_i$.

   (b) For $i \geq j + 1$ set $\widehat{e}_i = e_j$.

   [In this formulation $e_i - \widehat{e}_i$ stands for the candidate amount of bankers' wealth invested in deposits.]

4. Set $(k_i, w_i, R_{0i}) = (k(\widehat{e}_i), w(\widehat{e}_i), R_0(\widehat{e}_i))$ for each point i in the grid.

5. Consider the candidate $\{v_i\} = \{v(e_i)\}$. As an initial guess for $\{v(e_i)\}$, take some positive, non-decreasing function.

6. Identify, if it exists, the point in the grid m for which $v(e_m) \geq 1$ but $v(e_{m+1}) < 1$.

   (a) For $i < m + 1$, set $c_i = 0$.

   (b) For $i \geq m+1$, set $c_i = e_i - e_m$, and reset $(k_i, w_i, R_{0i}) = (k_m, w_m, R_{0m})$ and $\widehat{e}_i = \widehat{e}_m$.

   [In this formulation c_i stands for the candidate amount of bankers' wealth that active bankers consume. This procedure takes care of having $v(e_i) \geq 1$ for all e_i]

7. Use the following version of (13) to uniquely determine $R_{1i}^{1-\varepsilon}$ for each i :

$$(1-p_0)R_{1i}^{1-\varepsilon}-(1-p_1)R_{0i}-\frac{1}{\gamma}(p_0-p_1)[(1-\gamma)(1+r)-(1-\lambda)\frac{k_i}{k_i+w_i}]=0.$$

8. Use the following extended version of (14) to find $e_i^{1-\varepsilon}$ and e_i^{ε} for each i :

$$e_i^{1-\varepsilon}=\phi(1+r)w_i+(1-\psi)\{[(1-x_i)R_{0i+1}+x_iR_{1i+1}^{1-\varepsilon}]\widehat{e}_i+(1+r)(e_i-c_i-\widehat{e}_i)\},$$

$$e_i^{\varepsilon}=\phi(1+r)w_i+(1-\psi)[(1-x_i)R_{0i+1}\widehat{e}_i+(1+r)(e_i-c_i-\widehat{e}_i)].$$

9. Use the following version of (15) and (16) to check for a solution $x_i\in[0,1)$ for each i.

$$[(1-\varepsilon)v(e_i^{1-\varepsilon})+\varepsilon v(e_i^{\varepsilon})]R_{0i}-(1-\varepsilon)v(e_i^{1-\varepsilon})R_{1i}^{1-\varepsilon}\geq 0,$$

$$\{[(1-\varepsilon)v(e_i^{1-\varepsilon})+\varepsilon v(e_i^{\varepsilon})]R_{0i}-(1-\varepsilon)v(e_i^{1-\varepsilon})R_{1i}^{1-\varepsilon}\}x_i=0.$$

   If no such solution exists, set $x_i=1$, reset $(k_i,w_i,R_{0i})=(k(e_i-c_i),w(e_i-c_i),R_0(e_i-c_i))$ and check whether bankers could be interested in investing as bank capital some alternative $\widehat{\widehat{e}}_i\leq e_i-c_i$, keeping $e_i-c_i-\widehat{\widehat{e}}_i$ in a safe account.

   [This extends the notion of equilibrium in the main text to deal with equilibria in which all loans are systemic and active bankers may not necessarily reinvest all their wealth as bank capital.]

10. Use the following version of (5) to update the value function:

$$v(e_i)=\begin{cases}\psi+(1-\psi)\beta[(1-\varepsilon)v(e_{i+1}^{1-\varepsilon})+\varepsilon v(e_{i+1}^{\varepsilon})]R_{0i+1}, & \text{if } x_i\in[0,1),\\ \psi+(1-\psi)\beta(1-\varepsilon)v(e_{i+1}^{1-\varepsilon})R_{1i+1}^{1-\varepsilon}, & \text{if } x_i=1.\end{cases}$$

11. Check convergence, i.e. the proximity between the previous $\{v_i\}$ and the new $\{v(e_i)\}$. In case of convergence, save and report the solution, and finish. Otherwise, go to Step 5 and iterate again.

B Social welfare

Importantly, the patient agents who provide deposit funding to the banks at rate r just break-even in terms of their own net present value in all periods. Thus their net consumption flows as depositors make a zero net addition to W_t and we can safely leave them out of the expressions below. All other agents have a discount factor $\beta<1/(1+r)$, which is the one that we will use to discount the remaining consumption flows, including the negative flows associated with the taxes needed to cover the costs of deposit insurance when the systemic shock realizes (and the systemic bank goes bankrupt).[32] Focusing on the case in which bankers always reinvest 100% of their accumulated wealth as bank capital, social welfare at any period t, W_t, can be expressed as follows:

$$W_t = E_0 \left(\sum_{s=0}^{\infty} \beta^s \omega_{t+s} \right),$$

where

$$\omega_t = -e_t + [1 - \phi(1+\psi)]w_t + \beta\{y_{t+1} - (1+r)[d_t - \phi(1+\psi)w_t]\}, \tag{17}$$

$$y_{t+1} = gdp_{t+1} + (1 - \Delta_{t+1})k_t,$$

$$gdp_{t+1} = [(1-x_t)(1-p_0) + x_t(1-\varepsilon_{t+1})(1-p_1)]AF(k_t, 1), \tag{18}$$

$$\Delta_{t+1} = \delta + \{(1-x_t)p_0 + x_t[(1-\varepsilon_{t+1})p_1 + \varepsilon_{t+1}]\}(\lambda - \delta), \tag{19}$$

and $\varepsilon_{t+1} \in \{0, 1\}$ indicates whether the systemic shock realizes ($\varepsilon_{t+1} = 1$) or not ($\varepsilon_{t+1} = 0$) at the end of period t. In this decomposition, ω_t is the present value of the net consumption flows that the impatient agents derive from the activity of the economy in the production period between dates t and $t+1$. Such activity initially absorbs bank capital e_t from the bankers and pre-pays wages $[1 - \phi(1+\psi)]w_t$ to entrepreneurs and the workers who will not be bankers at $t+1$. It also pays wages ϕw_t to the active bankers and wages $\phi\psi w_t$ to the workers who will be bankers at $t+1$, but these two groups of impatient agents refrain from consuming and save those funds in the form of bank deposits. Finally, banks also advance the funds needed for firms to prepay physical capital at date t but, since those funds are precisely invested at date t, they bring about zero net consumption at date t.

At date $t+1$ (which explains the discount factor β in (17)), the impatient agents in the economy (including the taxpayers who pay, if needed, for the net costs associated with deposit insurance) appropriate (if positive) or contribute (if negative) the difference between gross output y_{t+1} and the gross repayments $(1+r)[d_t - \phi(1+\psi)w_t]$ to the patient agents who hold bank deposits. Gross output is the sum of GDP as conventionally defined, gdp_{t+1}, and depreciated physical capital, $(1 - \Delta_{t+1})k_t$. The expressions for these two components, (18) and (19), respectively, make clear that the GDP and the depreciation experienced by

[32] This assumption prevents the possibility of increasing our aggregative utilitarian measure of social welfare by using deposit insurance as an indirect means of redistributing wealth from the patient agents to the impatient agents.

physical capital at the end of a given period are affected both by the endogenous systemic risk-taking variable x_t and the realization of the systemic shock ε_{t+1}.

To further understand the sources of welfare in the expression above, notice that ω_t might be alternatively decomposed as the sum of net present value of the payoffs associated with the stakes held by each class of agents during the corresponding period :

1. Impatient agents other than active bankers and next-period bankers, who receive wages at t and consume: $+[1-\phi(1+\psi)]w_t$.
2. Patient agents who act as depositors, who break even in NPV terms: +0.
3. Entrepreneurs, who break even state-by-state: +0.
4. Tax payers, who pay deposit insurance costs at $t+1$: $-\beta[(1+r)d_t-(1-\lambda)k_t]x_t\varepsilon_{t+1}$.
5. Bankers as bank capital suppliers, who contribute e_t at t and receive bank equity returns at $t+1$:
$$-e_t+\beta[(1-x_t)R_{0t+1}+x_tR_{1t+1}]e_t.$$
6. Active and next-period bankers as suppliers of labor, who receive at $t+1$ the proceeds from having invested the wages earned at date t in bank deposits: $+\beta(1+r)\phi(1+\psi)w_t$.

Notice that this decomposition is not explicit about bankers' net consumption. Along a full-reinvestment path bankers only consume when they convert into workers again. Their consumption flow is implicit in the above components. Specifically, the total income inflow assigned to old and new bankers at $t+1$ in the expressions above are $[(1-x_t)R_{0t+1}+x_tR_{1t+1}]e_t+(1+r)\phi(1+\psi)w_t$, while the only income outflow assigned to them at that date is the equity capital e_{t+1} contributed to banks for next period of activity. Using (12), we obtain a net income flow of $\psi\{[(1-x_t)R_{0t+1}+x_tR_{1t+1}]e_t+(1+r)\phi w_t\}>0$ which indeed corresponds to the gross returns of the accumulated equity and the (deposited) last-period wages of the old bankers who convert into workers at date $t+1$ (which they entirely consume at that date).

All these expressions can be easily extended to the case in which, in certain periods, bankers voluntarily consume part of their wealth or keep part of it invested in deposits. All welfare computations in the quantitative part are based on the extended expressions, whose details we skip for brevity.

References

Acharya, Viral, and Tanju Yorulmazer (2007), "Too many to fail—An analysis of time-inconsistency in bank closure policies," *Journal of Financial Intermediation* 16, 1-31.

Acharya, Viral, and Tanju Yorulmazer (2008), "Cash-in-the-market pricing and optimal resolution of bank failures," *Review of Financial Studies* 21, 2705-2742.

Agénor, Pierre-Richard, Koray Alper, and Luiz A. Pereira da Silva (2009), "Capital requirements and business cycles with credit market imperfections," World Bank Policy Research WP No.5151.

Barro, Robert J. (2009), "Rare disasters, asset prices, and welfare costs," *American Economic Review* 99, 243-264.

Bernanke, Ben, Mark Gertler, and Simon Gilchrist (1999), "The financial accelerator in a quantitative business cycle framework,"in John Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Elsevier, Chapter 21, 1341-1393.

Bhattacharya, Sudipto, Arnoud W. A. Boot, and Anjan V. Thakor (1998), "The economics of bank regulation," *Journal of Money, Credit and Banking* 30, 745-770.

Boyd, John H., and Gianni De Nicoló (2005), "The theory of bank risk taking and competition revisited," *Journal of Finance* 60, 1329-1343.

Bruche, Max, and Javier Suarez (2010), "Deposit insurance and money market freezes," *Journal of Monetary Economics* 57, 45-61.

Brunnermeier, Markus K. and Yuliy Sannikov (2009), "A macroeconomic model with a financial sector," mimeo, Princeton University.

Carlstrom, Charles, and Timothy Fuerst (1997), "Agency costs, net worth and business fluctuations: A computable general equilibrium analysis," *American Economic Review* 87, 893-910.

Christiano, Lawrence, Roberto Motto, and Massimo Rostagno (2010), "Financial factors in economic fluctuations," ECB WP No. 1192.

Darracq Pariès, Matthieu, Christoffer Kok Sørensen, and Diego Rodriguez Palenzuela (2010), "Macroeconomic propagation under different regulatory regimes: evidence from an estimated DSGE model for the Euro Area," ECB WP No. 1251.

Dewatripont, Mathias, and Jean Tirole (1994), *The Prudential Regulation of Banks*, MIT Press.

Freixas, Xavier, and Jean Charles Rochet (2008), *Microeconomics of Banking*, 2nd edition, MIT Press.

Gertler, Mark, and Peter Karadi (2009), "A model of unconventional monetary policy," mimeo, New York University.

Gertler, Mark, and Nobuhiro Kiyotaki (2010), "Financial intermediation and credit policy in business cycle analysis," manuscript prepared for the *Handbook of Monetary Economics.*

Gertler, Mark, and Nobuhiro Kiyotaki, and Albert Queralto (2010), "Financial crises, bank risk exposure and government financial policy," mimeo, Princeton University

Hart, Oliver, and John Moore (1994), "A theory of debt based on the inalienability of human capital," *Quarterly Journal of Economics* 109, 841-879.

Hellmann, Thomas F., Kevin C. Murdock, Joseph E. Stiglitz (2000), "Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough?," *American Economic Review* 90, 147-165.

Holmstrom, Bengt, and Jean Tirole (1997), "Financial intermediation, loanable funds, and the real sector," *Quarterly Journal of Economics* 112, 663-691.

Jensen, Michael C., and William H. Meckling (1976), "Theory of the firm: Managerial behavior, agency costs and ownership structure," *Journal of Financial Economics* 3, 305-60.

Kareken, John H., and Neil Wallace (1978), "Deposit insurance and bank regulation: A partial–equilibrium exposition," *Journal of Business* 51, 413-438.

Kiyotaki, Nobuhiro, and John Moore (1997), "Credit cycles," *Journal of Political Economy* 105, 211-248.

Lorenzoni, Guido (2008), "Inefficient credit booms," *Review of Economics Studies* 75, 809-833.

Martinez-Miera, David, and Rafael Repullo (2010), "Does competition reduce the risk of bank failure?," *Review of Financial Studies* 23, 3638-3664.

Perotti, Enrico, and Javier Suarez (2002), "Last bank standing: What do I gain if you fail?," *European Economic Review* 46, 1599-1622.

Rancière, Romain, Aaron Tornell, and Frank Westermann (2008), "Systemic crises and growth," *Quarterly Journal of Economics* 123, 359-406

Repullo, Rafael (2004), "Capital requirements, market power, and risk-taking in banking," *Journal of Financial Intermediation* 13, 156-182.

Repullo, Rafael, and Javier Suarez (2000): "Entrepreneurial moral hazard and bank monitoring: A model of the credit channel," *European Economic Review* 44, 1931-1950.

Repullo, Rafael, and Javier Suarez (2004), "Loan pricing under Basel capital requirements," *Journal of Financial Intermediation* 13, 496-521.

Rietz, Thomas A. (1988), "The equity risk premium: A solution," *Journal of Monetary Economics* 22, 17-31.

Smets, Frank, and Rafael Wouters (2007), "Shocks and frictions in US business cycles: a Bayesian DSGE approach," *American Economic Review* 97, 586-606.

Stiglitz, Joseph E., and Andrew Weiss (1981), "Credit rationing in markets with imperfect information," *American Economic Review* 71, 393-410.

Van den Heuvel, Skander (2008), "The welfare costs of bank capital requirements," *Journal of Monetary Economics* 55, 298-320.