Banks' Endogenous Systemic Risk Taking*

David Martinez-Miera
Universidad Carlos III

Javier Suarez
CEMFI and CEPR

September 2014

Abstract

We develop a dynamic general equilibrium model that features endogenous systemic risk taking by banks. We use it to study the macro-prudential role of capital requirements. Bankers decide on the (unobservable) exposure of their banks to systemic shocks by balancing risk-shifting gains with the value of preserving their capital after such shocks. Capital requirements reduce systemic risk taking but at the cost of reducing credit and output in calm times, generating non-trivial welfare trade-offs. Interestingly, systemic risk taking is maximal after long periods of calm and may worsen if capital requirements are countercyclically adjusted.

Keywords: Capital requirements, Risk shifting, Credit cycles, Systemic risk, Financial crises, Macro-prudential policies.

JEL Classification: G21, G28, E44

* We would like to thank Matthieu Darracq Pariès, Giovanni Dell'Ariccia, Gianni De Nicolo, Mathias Dewatripont, Martin Ellison, Helmut Elsinger, Carlos González-Aguado, Hans Gersbach, Kebin Ma, Frederic Malherbe, Claudio Michelacci, Enrico Perotti, Rafael Repullo, Diego Rodriguez-Palenzuela, Ruben Segura, Philipp Schnabl, Nicholas Trachter, Alexandros Vardoulakis, Wolf Wagner, and the audiences at numerous conferences and seminars for helpful comments and suggestions. We acknowledge financial support from Bank of Spain and Spanish government grants ECO2010-17158 and JCI-2011-08963 (Martinez-Miera) and from the European Central Bank and Spanish government grant ECO2011-26308 (Suarez). Contact emails: dmmiera@emp.uc3m.es, suarez@cemfi.es.

1 Introduction

The deep and long lasting effects of the recent financial crisis have increased the motivation to better understand the contribution of banks to the generation of systemic risk. Systemic risk is a multifaceted phenomenon whose full understanding will require years of research. One of its facets consists of financial institutions being exposed to common shocks that, if sufficiently adverse, may take a significant fraction of them down at the same time and have a negative impact on the supply of credit to the real sector.[1]

In this paper we develop a model that explores the dynamic trade-offs underlying banks' decision to become exposed to rare but devastating common shocks. We model such decision as primarily influenced by the classical risk-shifting problem associated with leverage (Jensen and Meckling, 1976), which gets reinforced in the presence of explicit or implicit safety net guarantees (Kareken and Wallace, 1978). We analyze the extent to which capital requirements may contribute to reduce the resulting systemic risk taking and identify the trade-offs driving central issues in the discussions on the macro-prudential regulation of banks: the socially optimal level of the capital requirements and the extent to which such level should or not be adjusted over the credit cycle.

We consider banks owned by potentially long-lived *bankers* who are allowed to accumulate wealth by retaining past earnings.[2] Bankers' endogenously accumulated wealth is the only source of equity funding to banks, which banks need to be able to comply with the regulatory capital requirements.[3] Bank capital requirements influence bankers' incentives in regards to the adoption of systemic risk through two channels. First, the conventional leverage-reduction effect diminishes bankers' static gains from risk shifting. Second, capital requirements increase the demand for scarce bank capital in each state of the economy, rein-

[1] See, for example, Acharya (2011) or Hanson, Kashyap, and Stein (2011).

[2] This is like in other recent attempts to incorporate banks in dynamic general equilibrium setups, including Gertler and Kiyotaki (2010), Meh and Moran (2010), Gertler and Karadi (2011), Brunnermeier and Sannikov (2014), and He and Krishnamurthy (2014). Like in those papers, the analysis is simplified by making assumptions on heterogenous discounting and demographics (e.g. how agents switch roles in and out of banking) that prevent us from having to model the accumulation of wealth by agents other than bankers.

[3] Gertler et al. (2012) consider a setup where bankers' *inside* equity can be complemented with *outside* equity. However an agency problem limits the use of outside equity to a certain multiple of inside equity thereby preserving the essential properties of a model like ours, in which inside equity is the limiting factor.

forcing bankers' dynamic incentives to guarantee that their wealth (invested in bank capital) survives if a *systemic shock* occurs.[4]

Indeed, the loss of the capital devoted to systemic lending when a shock occurs allows the surviving bank capital to earn higher scarcity rents, producing a *last bank standing* effect similar to that identified in Perotti and Suarez (2002).[5] This effect reduces bankers' inclination towards systemic lending and gets reinforced when capital requirements are high (since they increase the relevant scarcity rents).[6] This last bank standing effect also helps explaining the key qualitative findings of the paper.

One of these findings is that systemic risk taking is maximal after several "calm periods" (i.e. periods in which the systemic shock does not occur), when output reaches its highest levels, bank equity is abundant, and the scarcity rents that it can appropriate diminish. Bankers react to the loss of shadow value of their wealth by increasing their appetite for systemic risk. This endogenously results in allocations where the vulnerability of the economy to systemic risk (i.e. the fraction of bank equity lost if the systemic shock occurs) is maximal precisely when credit supply and aggregate output are at their highest levels.

A second important finding is that strengthening capital requirements reduces the proportion of resources going into inefficient systemic investments, producing a lower loss of bank capital and a lower contraction in real activity when the systemic shock realizes. However, these gains come at the cost of reducing credit and output in calm times, generating an intuitive welfare trade-off. Measuring welfare as the expected present value of aggregate net consumption flows (since in our setup all agents are risk neutral), we find that there is a unique interior social welfare maximizing level of capital requirements.

A third qualitative implication due to the last bank standing effect is that making capital

[4]Our systemic shocks resemble the *rare economic disasters* considered in Rietz (1988) and Barro (2009), among others, which may empirically correspond to phenomena such as the bust of the US housing market around the summer of 2007. Rancière, Tornell, and Westermann (2008) develop a growth model in which levered firms make a choice between safe and risky growth strategies where the latter are exposed to this type of systemic shocks.

[5]In the imperfectly competitive setup explored by Perotti and Suarez (2002), banks are solely funded with deposits and the role of capital requirements is not discussed.

[6]As we further discuss in Section 6.3, in order for this mechanism to have the highest impact, it is convenient to resolve systemic crisis with the maximum dilution of the pre-existing equity of failed banks.

requirements cyclically adjusted is not necessarily welfare improving. Of course, reducing the capital requirement after a systemic shock would, ceteris paribus, reduce the credit crunch produced by the loss of bank capital. However, as bankers anticipate such countercyclical adjustment after a systemic shock, they also anticipate lower gains from protecting their capital against it and, thus, adopt higher systemic risk in the first place. We find that this negative ex ante effect may partly and even completely off-set the beneficial effect of reducing the credit crunch ex post.

To illustrate the quantitative implications of the model, we consider a parameterization in which social welfare turns out to be maximized under a relatively large capital requirement, 14%. To fix ideas, we compare the scenario with the optimal capital requirement with a baseline scenario with a 7% capital requirement (a level close to the requirements of core Tier 1 capital set by Basel III). We find that the unconditional mean of the fraction of bank equity devoted to support systemic lending under each of these requirements is 25% and 71%, respectively. The social welfare gain from having the optimal requirement rather than the low requirement is equivalent to a perpetual increase of 0.9% in aggregate net consumption—a large amount by macroeconomic standards. And the optimal capital requirement implies a much lower fall in aggregate net consumption, GDP, and bank credit in the year that follows a systemic shock.

Importantly, common macroeconomic aggregates such as GDP and bank credit have lower unconditional expected values under the optimal capital requirement than under the low requirement. This fall in average credit evidences that capital requirements improve the quality of credit at a cost in terms of the quantity of credit, and explains why it is not socially optimal to push capital requirements up to even higher levels (at which systemic risk taking might be reduced to zero but the implied credit level would be too low).

The model is suitable for the explicit analysis of the transition from a regime with a low capital requirement to another with a higher capital requirement. It allows to take into account the welfare losses implied by the credit crunch suffered when the requirements are raised but the economy has not yet accumulated the levels of bank capital that will characterize the new regime. In an illustration using our baseline parameterization, we find that, if starting from the low requirement regime and approaching some new target requirement in a linear way, it is socially optimal to implement the higher requirements over a number of years and to establish a more modest long-term target than if transitional costs were neglected.

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 the baseline parameterization and the main quantitative results. Section 6 shows the value of gradualism in the introduction of capital requirements, assesses the potential gains from making capital requirements cyclically adjusted, and contains several other extensions and discussions. Section 7 concludes. The appendices contain proofs, derive our measure of social welfare, and describe the numerical method used to solve for equilibrium.

2 Related literature

Our paper is related to recent efforts to understand the dynamic effects of banks on the real economy. Dynamic stochastic general 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 developed by Bernanke, Gertler, and Gilchrist (1999) but very few were explicit about banks.[7] Van den Heuvel (2008) undertakes the welfare analysis of capital requirements in a steady state environment in which bank deposits provide liquidity services to households, and banks are tempted to get involved in risk shifting.

The papers more closely related to our modeling of bank capital dynamics are Gertler and Kiyotaki (2010), Meh and Moran (2010), and Gertler and Karadi (2011), which also

[7]Some of the DSGE models attempting to capture banking frictions after the crisis adopt reduced-form approaches that do not include explicit foundations for regulation and, thus, impede a fully-fledged welfare analysis. See, for instance, Agénor et al. (2009), Christiano, Motto,and Rostagno (2013), Darracq Pariès, Kok Sorensen, and Rodriguez-Palenzuela (2011), and Gerali et al. (2010).

postulate a connection between bank capital and bankers' incentives.[8] These papers prescribe for bankers' wealth the same type of dynamics as for entrepreneurial net worth in Carlstrom and Fuerst (1997) and Kiyotaki and Moore (1997), among others. Similar capital dynamics also appears in Brunnermeier and Sannikov (2014), which captures a rich interaction between value-at-risk constraints, fire sales, and asset price volatility, and in He and Krishnamurthy (2014), which emphasizes the role of anticipating the disruption caused by states in which financial intermediaries hit occasionally binding financial constraints.[9] The main differences with respect to these papers is that our setup delivers an endogenous time-varying level of systemic risk-taking (and, associated with it, a time-varying bank failure rate) and that we focus the analysis on the macro-prudential role of bank capital requirements.

Our explicit focus on bank risk shifting and on how regulatory capital requirements interferes with it connects our contribution to long traditions in the corporate finance and banking literatures whose review exceeds the scope of this section. The seminal references on risk-shifting include Jensen and Meckling (1976) in a corporate finance context, and Stiglitz and Weiss (1981) in a credit market equilibrium context. Bhattacharya, Boot, and Thakor (1998) and Freixas and Rochet (2008) provide excellent surveys of subsequent contributions.

Risk shifting is identified by Kareken and Wallace (1978) as an important side effect of deposit insurance, and by Allen and Gale (2000) as the origin of credit booms and bubbles.[10] Banks' incentives to correlate their risk-taking strategies are justified by Acharya and Yorulmazer (2007) and Farhi and Tirole (2012) as a way to exploit the collective moral hazard problem that pushes the government to bail-out the banks when sufficiently many of them fail at the same time.

---

[8] In Gertler and Kiyotaki (2010) and Gertler and Karadi (2011), resembling Hart and Moore (1994), bankers have to partly finance their banks with their own wealth in order to commit not to divert the managed funds to themselves. Meh and Moran (2010) model market-imposed capital requirements along the same lines as Holmström and Tirole (1997), i.e. as a means to provide banks with incentives to monitor their borrowers.

[9] These two papers share with ours the analysis of the full non-linear solution of the corresponding model.

[10] When some relevant dimension of risk taking is unobservable, equilibrium risk taking may be excessive even without government guarantees. Yet the underpricing of those guarantees (or their flat pricing) may worsen the problem. Dewatripont and Tirole (1994) describe safety net guarantees as part of a social contract whereby depositors delegate the task of controlling banks' risk taking on the supervisory authorities who provide deposit insurance in exchange.

The role of capital requirements in ameliorating banks' risk shifting and their interaction with the incentives coming from banks' franchise values is a central theme in Hellmman, Murdock, and Stiglitz (2000) and Repullo (2004), where banks earn rents due to market power. Boyd and De Nicoló (2005) and Martinez-Miera and Repullo (2010) further explore this link in the presence of an additional entrepreneurial-incentive channel.

The dynamic incentives for prudence associated with the rise in the franchise value of surviving banks after a systemic crisis appear in Perotti and Suarez (2002) and Acharya and Yorulmazer (2008). However, differently from the prior tradition, the banks in our model are perfectly competitive and the relevant continuation value is attached to bank capital, which earns scarcity rents because bankers' endogenously accumulated wealth is limited.

3 The model

We consider a perfect competition, infinite horizon model in discrete time $t = 0, 1, ...$ in which all agents are risk neutral and production takes time and is subject to failure risk. To generate a role for banks, we assume that firms have to pay their factors of production in advance and banks are the sole providers of the required loans.[11] Banks are owned by some bankers who are the exclusive providers of bank equity, which in turn is needed to comply with a regulatory capital requirement. The next subsections describe and motivate each of these ingredients in detail.

3.1 Agents

The economy is populated by two classes of risk-neutral agents: *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 imposes a capital requirement to banks.

Patient agents have deep pockets. Their required expected rate of return is ρ per period, which can be interpreted as the exogenous return on some risk-free technology. Patient

[11] In subsection 3.2 we comment on a potential microfoundation of intermediation along the lines of Diamond (1984) and Holmström and Tirole (1997).

savers provide a perfectly elastic supply of funds to *banks* in the form of deposits but, due to unmodeled informational and agency frictions, cannot directly lend to the final borrowers.[12]

Impatient agents, of whom there is a continuum of measure one, are infinitely lived, have a discount factor $\beta < 1/(1+\rho)$, and inelastically supply a unit of labor per period at the prevailing wage rate w_t. Most impatient agents are mere *workers* and, as in other papers in the macroeconomic literature on financial frictions, we assume that *entrepreneurs* and *bankers* acquire their status in a random manner.[13] If the probability of a worker becoming a new entrepreneur (denoted η) or a new banker (denoted $\phi\psi/(1-\phi)$) in any given period are small enough, workers' impatience will imply that they do not accumulate any wealth prior to their change of status.[14] However, while remaining active entrepreneurs or bankers, financial frictions might motivate them to accumulate wealth.

To focus on bankers' dynamic incentives, we further assume that entrepreneurs' status is not persistent, so that they always develop their activity with zero wealth.[15] In contrast, we allow bankers to potentially remain active for several periods, accumulating wealth in the process via earnings retention. To start up such accumulation, we assume that they learn about their conversion into bankers one period in advance and, thus, can save the wage earned in such period in order to invest it as bank capital in the next one.[16] Finally, to prevent the population of active bankers (and their accumulated wealth) to grow without limit, we assume that bankers cease in their activity (and become pure workers again) with some time-independent probability ψ per period.[17]

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

[13] See, for example, Gertker and Kiyotaki (2010).

[14] We assume that impatient agents cannot borrow for pure consumption purposes. This could be due to the impossibility of pledging future income because of e.g. intertemporal anonymity. One could argue that banks can borrow from other agents and firms from banks because their end-of-period assets (loan to firms, depreciated physical capital, and net output) are *pledgeable*.

[15] Wealth accumulation by entrepreneurs or by mere workers will expand the number of state variables in the model, complicating the quantitative analysis.

[16] This is like in Bernanke and Gertler (1989). However, here bankers operate over potentially many periods and the bulk of their wealth dynamics in the parameterizations explored below is driven by the earnings retained while they are bankers.

[17] This probability ψ can be literarily interpreted as a retirement probability or, alternatively, as a reduced-form modeling of banks' payout policies or bankers' consumption decisions.

Prior assumptions produce stationary sizes η and ϕ for the populations of active entrepreneurs and bankers, respectively.[18]

3.2 Firms

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

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

units of the consumption good (which is the numeraire) at $t + 1$.[19] The binary random variable $z_{it+1} \in \{0, 1\}$, realized at $t + 1$, indicates whether the firm's production process succeeds ($z_{it+1} = 0$) or fails ($z_{it+1} = 1$). The parameters δ and $\lambda \geq \delta$ are the rates at which physical capital depreciates when the firm succeeds and when it fails, respectively.[20] The higher depreciation of capital in failed firms allows us to match the loss-given-default rates observed in corporate lending and makes firm failure in our model similar to a (firm specific) "capital quality shock" of the type explored in, e.g., Gertler and Kiyotaki (2010). Net output in case of success is the product of total factor productivity A and the function

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

with $\alpha \in (0, 1)$.[21] In case of failure, firms do not produce any output on top of depreciated capital.

The possible correlation of the failure shock z_{it+1} across firms is due to the exposure of firms to a common *systemic* shock $u_{t+1} \in \{0, 1\}$, whose bad realization $u_{t+1} = 1$ is assumed

[18]The size of the population of active entrepreneurs η is eventually irrelevant since, under the assumptions stated below, the technology exhibits constant returns to scale and firms' equilibrium profits are zero.

[19]Of course, *physical capital* (the good used as a production factor by firms) should not to be confounded with *bank capital* (the wealth that bankers contribute in the form of equity to the funding of the banks).

[20]In order to be able to summarize all the aggregate dynamics of the model through the evolution of a single state variable (bankers' wealth), we assume that physical capital can be transformed into the consumption good at all dates on a one-to-one basis.

[21]Notice that A is presented as a constant, so we abstract for simplicity from the type of productivity shocks emphasized in the real business cycle literature.

to occur with a constant independent small probability ε at the end of each period. The production technology can be operated in two modes that differ in their degree of exposure to the systemic shock: one is not exposed or *non-systemic* ($x_{it} = 0$), while the other is totally exposed or *systemic* ($x_{it} = 1$).

For firms operating in the non-systemic mode, z_{it+1} is independently and identically distributed across firms, and its distribution is independent of the realization of the systemic shock. Specifically, we have

$$\Pr[z_{it+1} = 1 \mid u_{t+1} = 0,\, x_{it} = 0] = \Pr[z_{it+1} = 1 \mid u_{t+1} = 1,\, x_{it} = 0] = \pi_0,$$

so, by the law of large numbers, the failure rate associated to any positive measure of non-systemic firms is constant and equal to π_0.

In contrast, we assume that all firms operating in the systemic mode have

$$\Pr[z_{it+1} = 1 \mid u_{t+1} = 0,\, x_{it} = 1\,] = \pi_1 < \Pr[z_{it+1} = 1 \mid u_{t+1} = 1,\, x_{it} = 1] = 1,$$

where failure in case of no shock ($u_{t+1} = 0$) is independently distributed across firms. Hence, the failure rate among systemic firms can be described as:

$$z_{t+1} = \begin{cases} \pi_1 & \text{if } u_{t+1} = 0, \\ 1 & \text{if } u_{t+1} = 1, \end{cases} \tag{3}$$

since systemic firms fail independently (with probability π_1) if the negative systemic shock does not occur, and simultaneously if it occurs.

Finally, following the *risk-shifting* literature, we assume that:

A1. $E(z_{it+1} \mid x_{it} = 1) = (1-\varepsilon)\pi_1 + \varepsilon > E(z_{it+1} \mid x_{it} = 0) = \pi_0$.

A2. $\pi_0 > \pi_1$.

Assumption A1 means that systemic firms are overall less efficient (i.e. yield lower total expected returns) than non-systemic ones. However, assumption A2 means that conditional on the systemic shock not occurring, non-systemic firms yield higher expected returns. This assumption implies that lending to systemic firms may be attractive to bankers protected by limited liability, who enjoy less defaults insofar as the systemic shock does not realize and suffer losses limited to their initial capital contributions otherwise.[22]

Entrepreneurs also run their firms under the protection of limited liability.[23] And to have a role for banks, 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} used at date t. The role for banks might be further justified along the lines of standard financial intermediation theory (e.g. Holmström and Tirole, 1997) by assuming that (i) entrepreneurs can unobservably undertake a third type of production process which is overall inviable but pays them high private benefits, and (ii) banks can operate some exclusive monitoring technology to prevent entrepreneurs from choosing such a process.[24]

The loan involves the *promise to repay* the amount $b_{it} \leq AF(k_{it}, n_{it}) + (1-\delta)k_{it}$ at $t+1$. This debt contract implies 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.[25] To capture bank competition we postulate that the tuple $(x_{it}, k_{it}, n_{it}, l_{it}, b_{it})$ is contractually set by each firm and its bank at date t in a manner that leaves any potential surplus with the firm subject to the participation constraint of the bank's owners.[26] Importantly, a firm's systemic orientation x_{it} is private information of the firm and its bank, which rules out regulations directly contingent on it.

3.3 Banks

Regulation obliges banks to finance at least a fraction γ_t of their one-period loans with equity capital i.e. with funds coming from bankers' accumulated wealth. Banks complement their funding with fully-insured one-period deposits taken from patient agents (as well as the

[22]It can be shown that with $\pi_1 \geq \pi_0$ no bank would get involved in the funding of systemic firms.

[23]Limited liability may be interpreted as an exogenous institutional constraint or an implication of anonymity, implying that entrepreneurs' contemporaneous or future wages cannot be used as collateral for entrepreneurial activities.

[24]Notice that the providers of labor and capital would not accept direct repayment promises from entrepreneurs because they would anticipate that, without bank monitoring, entrepreneurs would choose the inviable process.

[25]With non-negative loan rates and wages, we necessarily have $b_{it} \geq l_{it} = k_{it} + w_t n_{it} \geq k_{it} \geq (1-\lambda)k_{it}$.

[26]Nevertheless, as shown below, the constant returns-to-scale technology and the competitive product and factor markets make entrepreneurs' equilibrium profits equal to zero in all states. Meanwhile, the limited supply of bankers' wealth makes the appropriation of positive scarcity rents by bankers compatible with the competitive equilibrium.

bankers and would-be bankers who save their labor income until they can invest it in bank capital in the next date).[27] The deposit insurance scheme is paid for with contemporaneous non-distortionary taxes levied on impatient agents.[28]

We assume that banks hold *perfectly granular* loan portfolios, that is, extend infinitesimal loans to a continuum of firms, thus fully diversifying away firms' idiosyncratic failure risk.[29] Diversification, however, does not eliminate the systemic risk associated with lending to systemic firms. In fact, due to convexities induced by limited liability, bankers find it optimal to specialize their banks in either non-systemic or systemic loans.[30] Since banks are perfectly competitive and operate under constant returns to scale, we can refer w.l.o.g. to a representative *non-systemic bank* ($j = 0$) and a representative *systemic bank* ($j = 1$).

Each bank's balance sheet constraint imposes

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

for $j = 0, 1$, where l_{jt} denotes the loans made by bank j at date t, d_{jt} are its deposits, and e_{jt} is the equity provided by the bankers.[31]

The allocation of bank capital to each bank takes place in a perfectly competitive fashion. At any date t, bankers can invest their previously accumulated wealth as capital of the non-systemic bank, capital of the systemic bank or insured deposits; they can also consume all or part of their wealth.[32] If they contribute capital e_{jt} to bank j, they receive the free cash flow of the bank at $t + 1$ (i.e. the difference between payments from loans and payments to deposits) if it is positive, and zero otherwise. Bankers allocate their wealth based on their

[27] As it is well-known, deposit insurance reinforces banks' risk-taking incentives. However, in the absence of deposit insurance, systemic risk taking might still occur as result of a standard moral hazard problem, i.e. because banks' involvement in systemic lending is unobservable and occurs after deposits have been raised and priced.

[28] E.g. a tax on pure workers' consumption. Imposing this cost on impatient agents prevents the possibility of using deposit insurance as a means of redistribution of wealth from patient agents to impatient ones.

[29] We can think of this diversification as an easy-to-enforce regulatory imposition.

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

[31] Given that both classes of banks have access to unlimited deposit funding at a common rate, we can abstract from interbank lending and borrowing.

[32] Bankers can choose any mixture of these four options. They can, in particular, invest simultaneously in equity of the non-systemic and the systemic banks, although their risk-neutrality provides no special incentive for (or against) the diversification of their personal portfolios.

expectation about bank equity returns and the value of the resulting wealth across different possible states at $t+1$.

As it is standard in the analysis of corporations in a dynamic setup, banks take as given bankers' valuation of wealth across possible states at $t+1$, which provides the relevant stochastic discount factor for the valuation of securities held by bankers. Based on this and due to competitive pressure, banks formulate the participation constraint that guarantees that bankers are willing to provide the equity funding e_{jt} needed by each bank at t. As explained below, this constraint is taken into account when setting the terms of the lending contracts $(x_{it}, k_{it}, n_{it}, l_{it}, b_{it})$ with each of the entrepreneurs.

4 Equilibrium analysis

In our economy, bankers solve the genuinely dynamic optimization problems that determine how much of their wealth is invested as equity of the non-systemic bank e_{0t} or equity of the systemic bank e_{1t}. Banks instead are treated as perfectly competitive one-period ventures in which the bankers can invest. The fraction of total bank capital invested in systemic banks is denoted by $x_t \equiv e_{1t}/e_t \in [0,1]$.

We assume that banks play a pooling equilibrium in which the representative non-systemic bank optimizes on the terms of the contract signed with non-systemic firms, while the representative systemic bank prevents being identified as such (which would imply to be closed by the regulator) by mimicking the non-systemic bank in every aspect except the unobservable systemic orientation of its firms ($x_{it} = 1$). Importantly, in equilibrium, firms are indifferent between adopting a systemic or a non-systemic orientation because competitive factor and product markets, together with the constant returns to scale technology, imply that their equilibrium profits are zero.

Notice that when the systemic shock does not occur, the realized return on equity at the systemic bank (denoted R_{1t+1}) is higher than the return on equity at the non-systemic bank (denoted R_{0t+1}). This means that if these returns were observable one might ex post detect the systemic banks even in "calm times" (i.e. when the systemic shock does not realize).[33] However, we assume that bank accounts and managerial compensation practices are opaque enough to allow the owners of the systemic banks to appropriate the excess return without being discovered.[34]

4.1 Bankers' portfolio problem

Continuing bankers have the opportunity to reinvest the past returns of their wealth as bank capital for at least one more period. Let v_{t+1} denote the (stochastic) marginal value of one unit of an old banker's wealth at the time of receiving the returns from his past investment (right before learning whether he will remain active at $t+1$). If R_{jt+1} is the stochastic return paid by some security j at $t+1$, then an active banker's valuation of the security at date t will be $\beta E(v_{t+1}R_{jt+1})$, where βv_{t+1} plays the role of a stochastic discount factor.[35]

When a banker retires, which happens with probability ψ, his only alternatives are either to save the wealth as a bank deposit (earning a gross return $1+\rho$ at $t+1$) or to consume it (in which case one unit of wealth is worth just 1 at t). Given this agent's impatience and the small probability of ever becoming a banker (or entrepreneur) again, we assume that consuming is the optimal decision and, thus, the value of one unit of his wealth is just 1.[36]

With the prior point in mind and considering the optimization over the possible uses of one unit of wealth for a banker who remains active at $t+1$, we can establish the following Bellman equation for v_t:

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

The terms multiplied by $1-\psi$ reflect that the banker can optimize between the following

[33] A systemic bank is definitely detected if the systemic shock realizes, but at that point its capital is depleted and, under limited liability, there is no further punishment that can be imposed to its owners.

[34] The potential appropriability of the excess return from risk-shifting by bank managers might justify why the investment in bank equity is in the first place limited to the special class of agents that we call bankers, who might be interpreted as agents with the ability to either manage the banks or prevent being expropriated by their managers. This is consistent with the view in Diamond and Rajan (2000).

[35] This reflects that bankers' valuation of a unit of wealth may be different in different states of nature (e.g. depending on the scarcity of bankers' aggregate wealth). At an individual level, however, an old banker's wealth exhibits constant returns to scale, i.e. e_t units of wealth are worth $v_t e_t$.

[36] We check the validity of this assumption in all the parameterizations explored in the numerical part.

possibilities: (i) consuming the wealth, and (ii) investing in (a) deposits, (b) equity of the non-systemic bank, or (c) equity of the systemic bank.

Equation (5) implies a number of properties for v_t and the various possible equilibrium allocations of bankers' wealth. The possibility of consuming the wealth at t implies $v_t \geq 1$. Continuing bankers may decide to keep part of their wealth aside as bank deposits if $(1+\rho)E_t(v_{t+1}) \geq 1$ and the returns on bank equity (R_{0t+1} or R_{1t+1}) are small enough, i.e. $(1+\rho)E_t(v_{t+1}) \geq \max\{E_t(v_{t+1}R_{0t+1}), E_t(v_{t+1}R_{1t+1})\}$. However, in equilibrium, the last condition will never hold with strict inequality because in that case no banker would invest in bank capital and banks would not be able to give loans, which is incompatible with equilibrium under the technology described in (1).[37]

For brevity, the equilibrium conditions presented in the rest of the main text will focus on the case of *full reinvestment* in which $\beta \max\{E_t(v_{t+1}R_{0t+1}), E_t(v_{t+1}R_{1t+1})\} > \max\{1, \beta(1+\rho)E_t(v_{t+1})\}$. In this case, bankers' optimal portfolio decisions are to invest (i) only in equity of the non-systemic bank if $E_t(v_{t+1}R_{0t+1}) > E_t(v_{t+1}R_{1t+1})$, (ii) only in equity of the systemic bank if $E_t(v_{t+1}R_{1t+1}) > E_t(v_{t+1}R_{0t+1})$, or (iii) in any of the two if $E_t(v_{t+1}R_{0t+1}) = E_t(v_{t+1}R_{1t+1})$.

We will refer to $q_t \equiv \max\{E_t(v_{t+1}R_{0t+1}), E_t(v_{t+1}R_{1t+1})\}$ as bankers' required *value-weighted* return on wealth, which will be important in the analysis of the contract signed between firms and the non-systemic bank. To avoid problems interpreting the separating equilibrium that we characterize below, we will focus on parameterizations under which investing in the non-systemic bank is always sufficiently profitable to bankers, in which case $q_t = E_t(v_{t+1}R_{0t+1})$ for all t.[38]

[37]The Cobb-Douglas production technology and the Walrasian determination of equilibrium wages tends to make the marginal loan infinitely profitable when the supply of loans tends to zero, boosting the values of R_{0t+1} and R_{1t+1}.

[38]It is possible to analytically show that having a small measure of active bankers ($\phi \to 0$) or low risk-shifting incentives ($\pi_1 \to (\pi_0 - \varepsilon)/(1-\varepsilon)$) is sufficient to rule out equilibria with $x_t = 1$. Intuitively, with no entry of new bankers, if only a marginal unit of bankers' wealth survived a systemic shock, it would appropriate the going-to-infinity marginal returns to investment associated with the underlying production technology when the level of investment tends to zero. This would persuade some bankers to invest in equity of the non-systemic bank.

4.2 Lending contracts

This subsection describes how the representative non-systemic bank ($j = 0$) sets the terms of the contract that regulates the lending relationship with each of its funded firms. By definition, the non-systemic bank agrees on $x_{it} = j = 0$ with each of the firms that it finances. The representative systemic bank ($j = 1$) will simply mimic all the observable terms of this contract in order not to be detected and closed by the regulator.[39]

The non-systemic bank will set $(x_{it}, k_{it}, n_{it}, l_{it}, b_{it}) = (0, k_t, n_t, l_t, b_t)$, where k_t, n_t, l_t, and b_t solve the following problem:[40]

$$\begin{array}{cl} \max\limits_{(k_t,n_t,l_t,b_t,d_t,e_t)} & (1-\pi_0)[AF(k_t,n_t)+(1-\delta)k_t-b_t] \\ \text{s.t.} & E\{v_{t+1}[(1-\pi_0)b_t+\pi_0(1-\lambda)k_t-(1+\rho)d_t]\} \geq q_t e_t, \\ & l_t = k_t + w_t n_t,\ l_t = d_t + e_t,\ e_t \geq \gamma_t l_t. \end{array} \tag{6}$$

This problem maximizes the expected payoff of any of the funded entrepreneurs at the end of period t, subject to the constraints faced by the bank and the entrepreneur. When the firm does not fail, the entrepreneur obtains the difference between the gross output, $AF(k_t, n_t) + (1-\delta)k_t$, and the loan repayment, b_t. When the firm fails, he obtains zero.

The first constraint in (6) reflects bankers' participation constraint. The bank knows that an arbitrary stochastic payoff P_{t+1} offered in exchange for one unit of equity capital is acceptable to the bankers if and only if $E(v_{t+1}P_{t+1}) \geq q_t$, where v_{t+1} and q_t are taken as given. The payoffs that bankers receive at $t+1$ from the non-systemic bank are the gross repayments from the performing loans, $(1-\pi_0)b_t$, plus the payment coming from the recovery of depreciated physical capital in failed firms, $\pi_0(1-\lambda)k_t$, minus the payments due to depositors, $(1+\rho)d_t$. The last three constraints in problem (6) reflect: (i) the use of loans to pay firms' capital and labor in advance, (ii) the bank's balance sheet identity, and (iii) the regulatory capital requirement.

The fact that equity returns at the non-systemic bank are deterministic allows us to

[39] By definition, the systemic bank agrees on $x_{it} = j = 1$ with each of the firms that it finances.

[40] Since the constant returns-to-scale technology makes the optimal size of individual firms (and, hence, of individual loans) undetermined in equilibrium, it is useful to drop the firm subscripts i and to think of $(0, k_t, n_t, l_t, b_t)$ as the terms of a representative (linearly scalable) non-systemic loan.

divide both sides of the first constraint in (6) by $E(v_{t+1})$ and obtain

$$(1-\pi_0)b_t + \pi_0(1-\lambda)k_t - (1+\rho)d_t \geq R_{0t+1}e_t, \tag{7}$$

where R_{0t+1} is to be thought of the market-determined "required" return on equity at the non-systemic bank (that banks take as given).

In the problem stated in (6), the objective function is homogeneous of degree one and the constraints are such that, if some decision vector $(k_t, n_t, l_t, b_t, d_t, e_t)$ is feasible, then any multiple or fraction of such vector is also feasible. This implies that entrepreneurs' equilibrium payoff in the non-failure state (i.e. the term in square brackets in the objective function) will have to be zero.[41]

After expressing bankers' participation constraint like in (7), using the optimization conditions that emanate from (6), and the condition for labor market clearing, the following lemma establishes a number of relationships between some of the key endogenous variables of the model. The proof of the lemma is in Appendix A.

Lemma 1 *For a given expected return on equity at the non-systemic bank, R_{0t+1}, the optimal lending contract and the labor market clearing condition imply that, in equilibrium:*
(a) firms' aggregate demand for physical capital k_t satisfies

$$(1-\pi_0)[AF_k(k_t,1) + (1-\delta)] + \pi_0(1-\lambda) = (1-\gamma_t)(1+\rho) + \gamma_t R_{0t+1}, \tag{8}$$

(b) the market clearing wage rate w_t satisfies

$$(1-\pi_0)AF_n(k_t,1) = [(1-\gamma_t)(1+\rho) + \gamma_t R_{0t+1}]w_t, \tag{9}$$

(c) the minimal capital requirement is binding and the aggregate demand for equity capital e_t satisfies

$$e_t = \gamma_t(k_t + w_t), \ \ \textit{and} \tag{10}$$

(d) the gross loan rate $1 + r_t = b_t/l_t$, satisfies

$$1 + r_t = \frac{1}{1-\pi_0}\{[(1-\gamma_t)(1+\rho) + \gamma_t R_{0t+1}] - \pi_0(1-\lambda)\frac{k_t}{k_t + w_t}\}. \tag{11}$$

[41] This is conclusion follows from standard reasoning under perfect competition and constant returns to scale: if the referred payoff were strictly positive, entrepreneurs would like to scale their firms up to infinity; if it were strictly negative, they would simply not operate their firms.

Equations (8) and (9) are a natural extension of the conditions associated with the canonical problem of perfectly-competitive firms in static production theory. These equations take into account several features of the extended problem. First, the production process is intertemporal and subject to failure risk. Second, expected gross output at $t+1$ is partly net output and partly depreciated capital. Third, the factors k_t and n_t are pre-paid at t using bank loans and, hence, their effective cost is affected by the bank's weighted average cost of funds, which is $(1-\gamma_t)(1+\rho)+\gamma_t R_0$ because the capital requirement $e_t \geq \gamma_t l_t$ is always binding.[42]

Bank frictions affect the real sector through the cost of the loans that firms use to finance their factors of production. For given capital requirement γ_t, increasing the required rate of return on bank capital R_{0t+1} increases the competitive bank loan rate, pushing firms to reduce their scale, which, after taking labor market clearing into account, implies that both k_t by (8) and, recursively, w_t by (9) fall.[43] Hence, the demand for bank capital described in (10) is decreasing in R_{0t+1}. With these ingredients, determining the equilibrium path for R_{0t+1} will result from adding the supply side of the market for bank capital and making sure that such market clears at each date.

4.3 The dynamics of the supply of bank capital

For the purposes of this subsection, let us think of e_{t+1} as the aggregate supply of bank capital at date $t+1$. Along a full reinvestment path, e_{t+1} coincides with the total wealth of active bankers at the beginning of period $t+1$, which is made up of two components: (i) the gross return of the labor income, ϕw_t, that bankers invested in bank deposits at date t (to be able to invest it in bank equity at $t+1$), and (ii) the gross returns on the wealth that continuing bankers invested in bank capital at date t, $(1-\psi)e_t$.[44] This results in the

[42] The minimal capital requirement is binding because the bank finds insured deposit funding cheaper than the equity funding coming from its owners' scarce wealth. (Notice that the bankers could always invest their wealth as insured deposits, so we must have $R_0 \geq 1+\rho$.)

[43] The same effects follow from an increase in γ_t, for given $R_{0t+1} > 1+\rho$.

[44] Appendix B states equilibrium conditions for the general case in which active bankers may find it optimal to consume part of their accumulated wealth or to keep part of it inverted as bank deposits. In the numerical solution we also check for the optimality of bankers and would-be bankers to invest their labor income in deposits for one period.

following law of motion for e_{t+1}:

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

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

From the point of view of date t, R_{0t+1} is deterministic while R_{1t+1} is a random variable that solely depends on the realization of u_{t+1}. When needed, we will use superindeces 0 and 1 to identify the ex-post value conditional on $u_{t+1} = 0$ and $u_{t+1} = 1$, respectively, of those variables that vary with the shock. If the systemic shock does not realize, one unit of capital of the systemic bank yields the gross return

$$R^0_{1t+1} = \frac{1-\pi_1}{1-\pi_0} R_{0t+1} + \frac{1}{\gamma_t}\frac{\pi_0-\pi_1}{1-\pi_0}[(1-\gamma_t)(1+\rho) - (1-\lambda)\frac{k_t}{k_t+w_t}], \tag{13}$$

which is larger than R_{0t+1} under assumption A2. This expression is found taking into account that the systemic bank mimics the non-systemic bank in every decision but, when the systemic shock does not realize, the default rate on its loans is π_1 rather than π_0.

In contrast, under most reasonable parameterizations, if the systemic shock realizes, the systemic bank becomes insolvent and, by limited liability, its owners realize a gross equity return $R^1_{1t+1} = 0 < R_{0t+1}$.[45] In this case, the aggregate bank capital available at date $t+1$ can be described as a random variable with the following law of motion:

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

driven by the realization of the aggregate shock u_{t+1}.

Before closing this subsection, it is convenient to look back at (5) and use (14) to summarize the conditions for the compatibility of particular values of x_t with bankers' optimal portfolio decisions.

[45] A sufficient condition for the systemic bank to fail when the systemic shock realizes is that the capital requirement γ_t is lower than the rate of depreciation of physical capital in failed projects λ. The condition $\gamma_t < \lambda$ holds in all the quantitative analysis below—even when γ_t is set at its social welfare maximizing value.

Lemma 2 *Bankers' optimization in an equilibrium with* $x_t \in [0, 1)$ *requires:*

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

Moreover, if (15) holds with strict inequality, the equilibrium must involve $x_t = 0$.

The corner solution without systemic risk taking ($x_t = 0$) that emerges when (15) holds with strict inequality will be formally captured when solving for equilibrium by imposing the complementary slackness condition:

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

4.4 Equilibrium

In any full-reinvestment equilibrium, the state of the economy at any date t can be summarized by a single state variable: the total wealth available to the active bankers e_t. As described in (14), e_t is determined by, among other factors, the realization of the systemic shock u_t at the end of the prior period. The equilibrium values of all other variables can be expressed as functions of the state variable e_t that satisfy the relevant individual optimization and market clearing conditions (already established in previous sections). More formally:

Definition 1 *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* $(v(e), x(e), k(e), w(e),\ R_0(e), R^0_1(e))$ *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* $\{v_t, x_t, k_t, w_t, R_{0t+1}, R^0_{1t+1}\}_{t=0,1,...}$ *that they generate satisfy:*

1. *Optimization by all the relevant agents.*

2. *The clearing of all markets.*

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

Thus, along an equilibrium path, the equilibrium values of the marginal value of bank capital v_t, the fraction of bank capital allocated to the systemic bank x_t, the physical capital used by firms k_t, the wage rate w_t, the return on equity at the non-systemic bank R_{0t+1}, and the return on equity at the systemic bank when the systemic shock does not occur R^0_{1t+1} can be found by evaluating the various components of the tuple $(v(e), x(e), k(e), w(e), R_0(e), R^0_1(e))$ at the amount of aggregate bank capital $e = e_t$ available at date t. And the amount of bank capital available in the subsequent period can be found by feeding (14) with these variables and the corresponding realization u_{t+1} of the systemic shock at $t + 1$.

Appendix B describes the numerical solution method used to solve for equilibrium. The appendix relaxes requirement 3 in the above definition to allow for solutions in which, in some states, bankers optimally devote part of their wealth to consume or to invest in bank deposits.

4.5 The last bank standing effect

Given the fixed supply of labor and the underlying constant-returns-to-scale technology, the aggregate returns to bank lending in our economy are marginally decreasing. This makes a marginal unit of bankers' wealth (the key resource needed to expand banks' lending capacity) more valuable when bankers' aggregate wealth is more scarce.

Intuitively, increasing e expands banks' lending capacity, makes loans cheaper, and allows firms to expand their activity, which in equilibrium, after wages adjust, implies devoting more physical capital to production. But then, like in the neoclassical growth model, the fixed supply of labor makes the aggregate return on physical capital marginally decreasing. Consequently, the marginal value of bank lending and the scarcity rents appropriated by bank capital, reflected in $v(e)$, also decrease with e.[46]

The decreasing marginal value of bank capital, in combination with the dynamics of bank capital described in (14), implies that after sufficiently many periods without suffering a systemic shock, the economy converges to what we denote as its *pseudo-steady state* (PSS): a state in which all aggregate variables remain constant insofar as the systemic shock does

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

not realize. If such shock realizes, the fraction of e invested by bankers in equity of the systemic bank is lost and the process of accumulation of bankers' wealth re-starts.

To understand the intuition driving bankers' systemic risk-taking decisions, notice that, as shown in equation (13), conditional on not suffering the systemic shock, systemic lending is more profitable than non-systemic lending, $R^0_{1t+1} > R_{0t+1}$. So satisfying the non-arbitrage condition (15) requires a sufficiently high valuation for the equity that survives the systemic shock, $v(e^1_{t+1})$. Since $v(e)$ is decreasing, this in turn requires that a sufficiently low amount of surviving equity, e^1_{t+1}. Intuitively, the bankers who give up the gains from risk-shifting must be compensated by the expectation of obtaining a large revaluation of their surviving wealth when the systemic shock occurs. In other words, banks' systemic risk taking incentives need to be compensated by a sufficiently strong last bank standing effect.[47]

By the law of motion described in (14), larger aggregate systemic risk taking x_t implies, other things equal, a larger aggregate loss of bank capital when the shock occurs, and hence a lower e^1_{t+1} and a larger $v(e^1_{t+1})$. This establishes a self-equilibrating mechanism in the operation of the last bank standing effect and leads to the existence of a unique x_t that solves equations (15) and (16).

4.6 Social welfare

A natural measure of social welfare W_t in this economy is the expected present value of the aggregate net consumption flows of the various agents from date t onwards. This measure can be obtained from different decompositions. One can infer the net consumption flow that the economy generates for each class of agents in each date t and aggregate across agents. Equivalently, one can just look at the differences between the aggregate quantity of the consumption good the economy produces at the end of a period and the quantity which is reutilized as a factor of production (physical capital) in the next period. Appendix

[47] This last bank standing effect resembles the traditional *charter value effect* of the microeconomic banking literature but is different along several important dimensions. First, it has its roots in a general equilibrium effect (the temporary impact of the systemic shock on the scarcity of bank capital) rather than in some permanent rents due to imperfect competition. Second, as shown below, our last bank standing effect gets reinforced when γ increases, whereas the usual charter value effect may be weaken by the negative impact of capital requirements on bank profits (see, for instance, Hellmman, Murdock, and Stiglitz, 2000).

C provides an explicit expression for W_t (equation (31)) and an associated flow measure of welfare nc_t, which are explained there using these two intuitive decompositions.

5 Numerical results

Our baseline quantitative results are obtained under a time-invariant capital requirement $\gamma_t = \gamma$ for all t. For illustration purposes, we will compare the results obtained with a reference capital requirement of 7% (γ=0.07), close to the overall level of the Basel III core equity requirement, with those obtained with the requirement of 14% (γ=0.14) that, under the parameterization presented below, maximizes the *unconditional* expected value of W_t.[48]

In Section 6, we analyze some cases with time-varying or state contingent capital requirements. In particular, we assess the implications of moving from a regime with γ=0.07 to a regime with higher capital requirements in a gradual way. We also assess potential gains from giving a pro-cyclical or counter-cyclical profile to γ_t.

5.1 Baseline parameterization

Our quantitative results are based on assuming that one period in the model corresponds to one year in calendar time. Table 1 contains the parameterization chosen to illustrate the quantitative properties of the model. Section 6 analyzes the sensitivity of the results to changes in some of the parameters.

The model is quite parsimonious: it has the 11 parameters listed in Table 1 (plus the capital requirement γ, if taken as given) and a single binary i.i.d. aggregate shock (the systemic shock). The discount rate of the patient agents, ρ, is chosen equal to 2% to capture a situation with low real interest rates such as that observed in the years leading to the 2007 financial crisis. Consistent with the literature on external financing frictions, the discount rate of the impatient agents is set approximately twice as large as ρ.[49] The total factor

[48] To compute the unconditional expected value of W_t we calculate a weighted average of W_t over the points of the ergodic distribution of e_t, with the weights given by the relative frequency with which those points are visited in a simulated path of 50,000 periods.

[49] For instance, in Iacoviello (2005) the spread between the discount rate of the borrowing entrepreneurs and that of the patient households that finance them is 4%. In Carlstrom and Fuerst (1997) and Gomes,

productivity parameter A, which only affects the scale of the variables in levels, is set equal to 2 (which conveniently produces macroeconomic aggregates with one or two digits in levels, making them just easier to report).

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

Patient agents' discount rate	ρ	0.02
Impatient agents' discount factor	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	π_0	0.03
Idiosyncratic default rate of systemic firms	π_1	0.018
Probability of a systemic shock	ε	0.03
Bankers' payout rate	ψ	0.20
Fraction of wages devoted to new bank capital formation	ϕ	0.05

Following convention, the elasticity of physical capital in the production function α is fixed so as to produce a share of labor income in GDP of about 70%, and the depreciation rate of physical capital in successful firms δ is chosen so as to match an aggregate physical capital to GDP ratio in the range of 3 to 4. The depreciation of physical capital in failing firms λ is consistent with 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.[50]

The default probabilities π_0 and π_1, and the systemic shock probability ε are set so as to have sufficient potential room for risk shifting and for significant aggregate losses due to it. Specifically, the current choices are compatible with the conditions $(1-\varepsilon)\pi_1 + \varepsilon > \pi_0 > \pi_1$ established in assumptions A1 and A2, and imply unconditional *expected* default rates in the range from 3% (if all firms are non-systemic) to 4.7% (if all firms are systemic). The probability of a systemic shock is set at 3%, so that a systemic crisis occurs on average once every 33 years.

Yaron, and Zhang (2003), the spread is 5.6%.

[50] To explain why a depreciation rate of $\lambda = 0.35$ of physical capital produces an overall LGD of 45%, notice that the loans in this model also finance firms' wages.

Bankers' exit rate ψ is set at 0.20, which strictly speaking implies that bankers have an average active life of 5 years over which to accumulate wealth by retaining all their earnings. Perhaps more realistically, $\psi = 0.20$ can also be interpreted as a situation in which 80% of the aggregate wealth resulting from bankers' prior activity remains reinvested as bank equity (rather than being *paid out*) in every period. Finally, setting parameter ϕ equal to 0.05 means that additions to active bankers' available wealth coming from both continuing and new bankers is roughly equal to 5% of aggregate labor income per period.

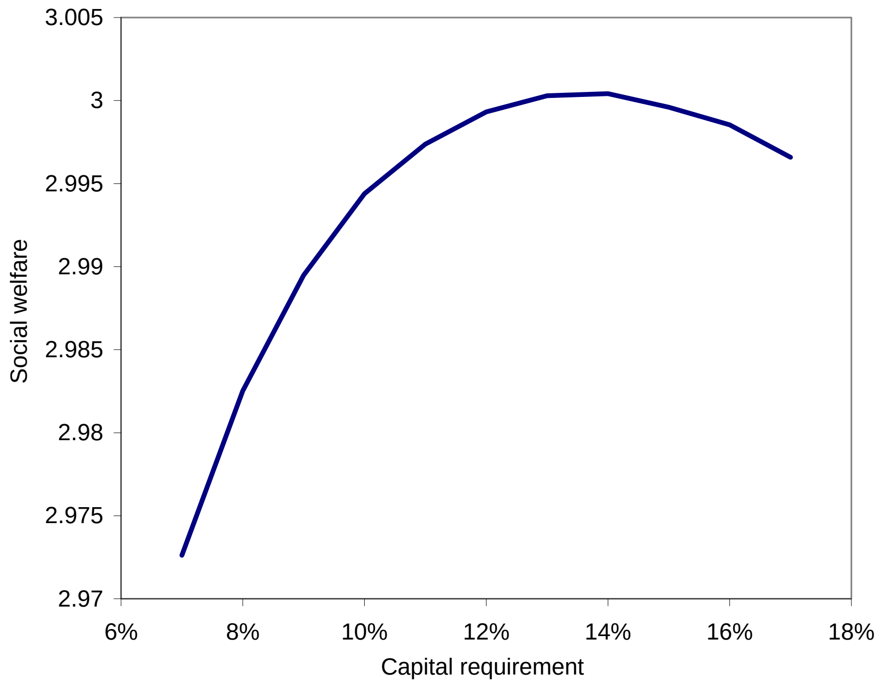

Figure 1 Social welfare, $(1-\beta)E(W_t)$, as a function of the capital requirement γ

5.2 Graphical presentation of the results

The method used to solve the model relies on value function iteration to obtain $v(e)$. As described in Appendix B, all equilibrium conditions are solved in their full, potentially non-linear form.[51] No problem of multiplicity of equilibria has been detected.

[51] Linearization is only used locally, for interpolation purposes, when the value function has to be evaluated at values of e not included in the initial grid.

Figure 1 is generated by solving the model for a grid of values of γ and by computing the unconditional expected value of social welfare, $E(W_t)$, under each of them. The figure describes welfare as the *certainty-equivalent consumption flow* $(1-\beta)E(W_t)$ which, if received as a perpetuity, would have a present discounted value of $E(W_t)$.[52]

Figure 2 depicts, for the optimal capital requirement of 14% (γ=0.14) and for the illustrative alternative value of 7% (γ=0.07), the functions that describe the marginal value of one unit of bank capital $v(e)$ (top panel) and the fraction of bank capital devoted to make systemic loans $x(e)$ (bottom panel). Both functions are depicted for a range of values of e that includes the ranges relevant under each of the compared capital requirements.

This figure evidences that the greater scarcity of bank capital induced by a higher capital requirement implies a higher marginal value of capital $v(e)$ at every level of capital e. More importantly, systemic risk taking is non-decreasing in bankers' aggregate wealth e (in fact, strictly increasing whenever $x(e) > 0$ and bankers fully reinvest their wealth as bank equity) and is lower at every e with the optimal capital requirement than with the low requirement.

The effect of the capital requirement on systemic risk taking is partly explained by its impact on the last bank standing effect. A higher γ implies higher scarcity and, thus, a higher equilibrium value of bank capital, $v(e)$, which increases bankers' incentives to guarantee that their wealth survives a systemic shock.

[52] The use of impatient agents' discount factor β in the discounting of the relevant consumption flows is justified in Appendix C.

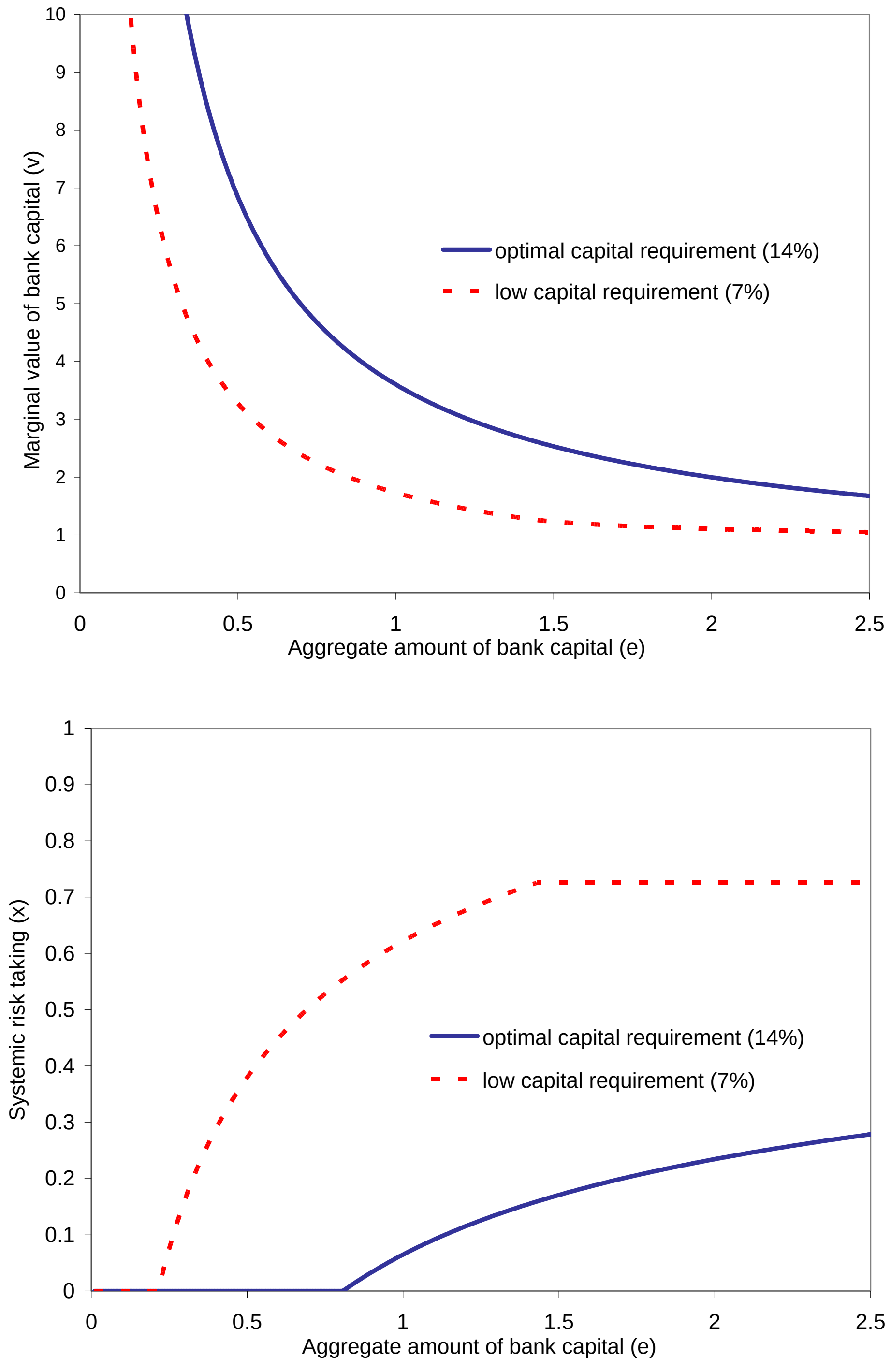

Figure 2 $v(e)$ and $x(e)$ under low and optimal capital requirements

The interaction between bankers' systemic risk taking and the endogenous dynamics of bank capital can be further explained by looking at Figure 3. The panels on the left correspond to the economy with $\gamma = 0.07$ and those on the right to the economy with $\gamma = 0.14$. The solid curve in each of the top panels represents the phase diagram mapping the amount of bank capital in one period e_t onto the amount available in the next period if the systemic shock does not occur e^0_{t+1}. This schedule is strictly increasing except if e_t is large enough for continuing bankers to consume part of it as a voluntary dividend, in which case the schedule would become flat. The schedule is strictly concave while banks keep reinvesting all their wealth as bank equity and becomes linear (with a slope of $(1-\psi)(1+\rho)$) at the point they save part of their wealth as deposits (an option explicitly considered in Appendix B)

The dashed curve in each of the top panels represents the mapping from e_t onto the capital available to bankers in the next date if the systemic shock occurs, e^1_{t+1}. The vertical distances between the solid and the dashed curves measure the loss of bank capital when the economy is hit by the systemic shock. The loss is larger not only in absolute but also in relative terms for higher values of e because the fraction of equity invested in the systemic bank increases with e (recall Figure 2).

For sufficiently low values of e, we get $x(e) = 0$, in which case the two curves merge ($e^0_{t+1} = e^1_{t+1}$). For sufficiently large values of e, the dashed curve may also become flat, if bankers start consuming part of their wealth, or linearly increasing (with slope $(1-\psi)(1+\rho)$) if they start investing some of their wealth in deposits).

The point where the solid phase diagram in each of the top panels intersects the 45-degree line identifies the corresponding pseudo-steady state. Interestingly, the PSS value of e is the highest point in the ergodic support and, thus, is associated with the highest level of systemic risk taking (since $x'(e) > 0$). This is also the point where the realization of the systemic shock implies the largest loss of bank capital and the largest subsequent contraction of credit.

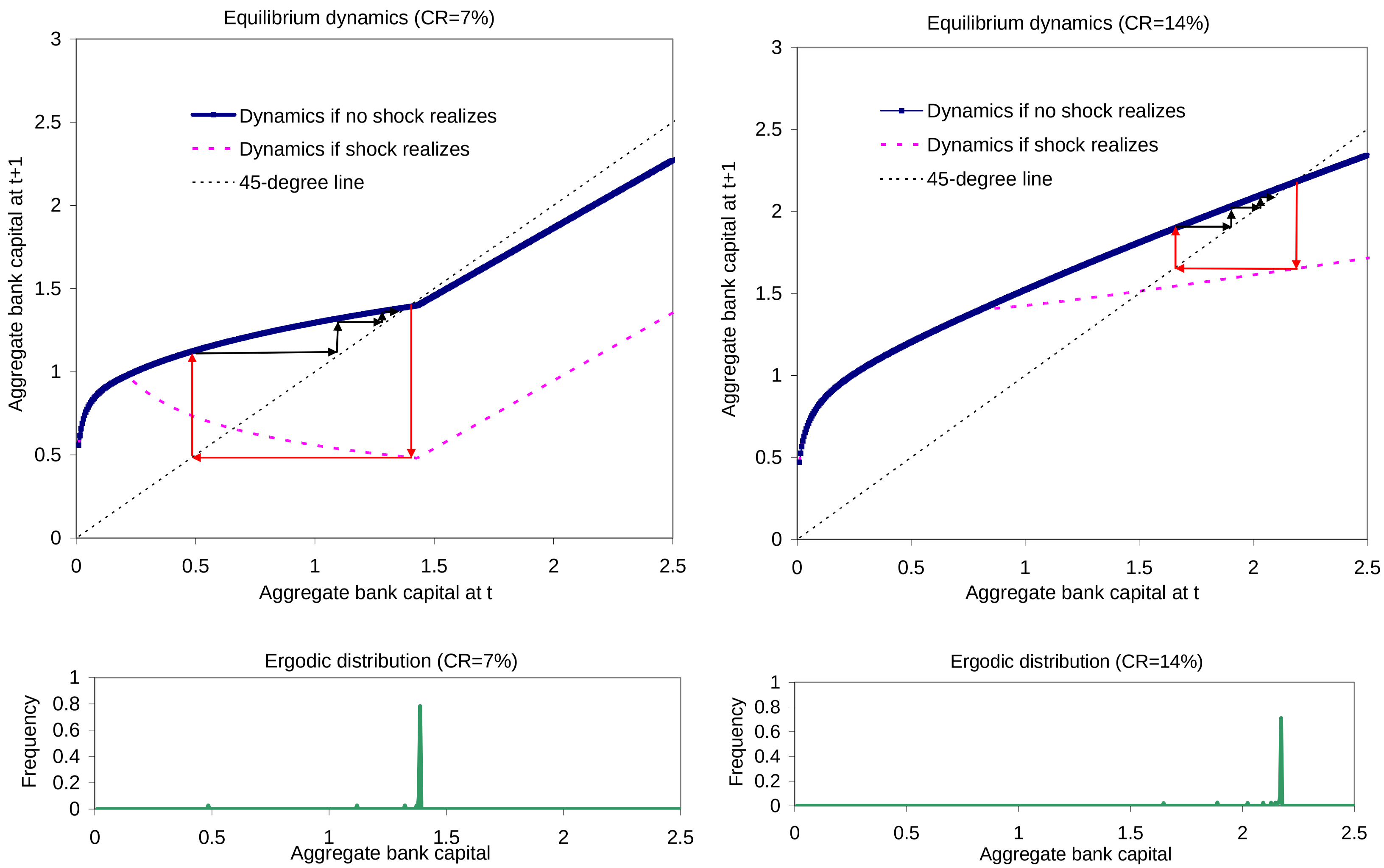

Figure 3 Equilibrium dynamics with low (CR=7%) and optimal (CR=14%) capital requirements

The arrows on each panel identify the path of crisis and recovery for the (most frequent) situation in which the economy fully returns to its PSS without suffering a second systemic shock.[53] With $\gamma = 0.07$ ($\gamma = 0.14$) the simulated economy fully recovers in a minimum of 6 (8) years.[54]

The bottom panels in Figure 3 depict the relative frequencies with which different values of e are visited along sufficiently long histories of the economy. Consistently with the aggregate shock being so infrequent, our economy spends most of the time (about 80% or more) in the pseudo-steady state. Other points with positive frequencies are those visited along recovery paths followed after suffering a systemic shock at the PSS or elsewhere in the recovery from a previous systemic shock.

Figures 2 and 3 reflect some of the considerations relevant for the welfare comparison between the economies with $\gamma = 0.07$ and $\gamma = 0.14$. The unresolved risk-shifting problem reflected in $x(e) > 0$ produces "static" losses due to the inefficiency of operating the production technology in a way that implies a larger unconditional failure rate. Risk shifting also produces "dynamic" losses due the reduction in bank equity (and the subsequent reduction in banks' lending capacity) that follows the realization of a systemic shock, causing amplification and propagation of its effects.

Rising the capital requirement reduces systemic risk taking but does not guarantee a larger social welfare. This is because of the negative impact on banks' lending capacity. For instance, in Figure 3, the PSS level of bank equity is not twice as large with $\gamma = 0.14$ as with $\gamma = 0.07$, meaning that the optimal capital requirement cannot sustain as much credit (and economic activity) as the lower capital requirement. This is a key factor explaining why the optimal capital requirement is not the one which pushes $x(e)$ all the way down to zero.

5.3 Quantitative details

Table 2 reports the unconditional expected values of the main endogenous variables of the model under the two levels of the capital requirement that we compare. The first variable in

[53] Recall that in our calibration systemic shocks are i.i.d. and occur with a probability of 3% per year.

[54] The differences between the two economies in this respect are hard to see in the figures since in the last periods prior to returning to PSS, the economy is very close to it.

the list is social welfare, reported in certainty equivalent consumption terms, like in Figure 1. We find that the difference between the welfare associated with $\gamma = 0.14$ and that associated with $\gamma = 0.07$ is equivalent to a perpetual increase of 0.9% in aggregate net consumption. The main reason for this gain is the lower average fraction of bank capital devoted to support systemic lending: 25% rather than 71%.

Table 2
Main endogenous variables
(Unconditional expected values and percentage difference across regimes)

	$\gamma = 7\%$	$\gamma = 14\%$	% diff.
Welfare* (=equivalent perpetual consumption flow)	2.973	3.000	0.93
GDP*	4.539	4.154	-8.50
Bank credit (l)	19.30	15.28	-20.84
Bank equity (e)	1.35	2.14	58.31
Loan rate (r) (in %)**	4.1	5.6	1.5
Deposit insurance costs*	0.159	0.038	-76.21
Value of one unit of bank capital (v)	1.37	1.90	38.13
Fraction of capital invested in systemic banks (x)	0.705	0.248	-64.85

*See Appendix C for exact definitions of these variables. **Difference reported in percentage points.

Table 2 also shows that with the optimal capital requirement bank loans are on average more expensive (with an unconditional mean loan rate of 5.6%) than with the low requirement (4.1%). This implies obtaining significantly lower values in macroeconomic aggregates such as GDP and bank credit, whose unconditional means are, respectively, 8.5% and 20.8% lower with the optimal requirement than with the low requirement.[55]

When the systemic shock occurs, the loss in bank capital reduces credit and investment for a number of periods, until the economy returns to its pseudo-steady state. Table 3 reports the fall in the variables listed in Table 2 that occur when the economy gets hit at its PSS by a systemic shock.

[55] The analysis evidences that bank credit and GDP are bad proxies of social welfare in the presence of systemic risk. These variables do not properly reflect the relatively low social net present value of the marginal investments undertaken when systemic risk taking is large. GDP in particular does not capture the sizeable losses (e.g. in physical capital not recovered from failed projects) incurred when systemic risk materializes.

Table 3
Percentage change in main variables after a systemic shock
(In the period after the shock, relative to PSS value if no shock occurs)

	$\gamma = 7\%$	$\gamma = 14\%$
Aggregate net consumption*	-17.49	-4.60
GDP*	-33.68	-10.03
Bank credit (l)	-65.83	-24.37
Bank equity (e)	-65.83	-24.37
Loan rate (r) (difference in percentage points)	11.8	2.6
Value of one unit of bank capital (v)	164.28	25.43
Fraction of capital invested in systemic banks (x)	-49.98	-23.66

*See Appendix C for exact definitions of these variables.

This table highlights the important cliff effects that the model can generate and how they depend on the level of capital requirements. When the systemic shock occurs it wipes out a fraction of bank capital and leads to a contraction in the supply of credit. Under the low capital requirement, loan rates increase by 11.8 percentage points following a systemic shock, while they only increase by 2.6 percentage points under the optimal capital requirement. Aggregate net consumption, GDP, and bank credit fall by 17%, 34%, and 66% with the low requirement and by only 5%, 10%, and 24% with the optimal requirement.

Table 4 describes the values of several other macroeconomic and financial ratios across the two compared economies. Some ratios simply reflect some of our targets when the choosing the model parameters. Other ratios constitute more genuine results and point to weak capital regulation (low γ) as a potential cause of *financial exuberance.* In particular, the ratio of credit to GDP (presented in Basel III as a candidate benchmark for macroprudential policy) happens to be higher in the economy with low capital requirements, i.e. when the endogenous level of systemic risk taking is higher. This is consistent with the perceptions of the early proponents of the macroprudential approach to bank regulation, which identified insufficient regulation as a cause of excessive credit (e.g. Borio, 2003).

Table 4
Other macroeconomic and financial ratios
(Unconditional expected values of ratios; percentages when indicated)

	$\gamma = 7\%$	$\gamma = 14\%$
Labor income/GDP*	0.67	0.67
Physical capital/GDP	3.58	3.01
Bank credit/GDP	4.25	3.68
Deposit insurance costs/GDP (ratio of expected values, %)	3.5	0.9
Net return on equity (ROE) at non-systemic banks (%)	10.2	16.7
Net ROE at systemic banks if systemic shock does not realize (%)	18.7	21.2
Banks' payout/Bank capital (%)	21.2	22.4

*See Appendix C for an exact definition of GDP.

The results in prior tables suggest that the quantitative implications of capital requirements can be quite sizeable. They also suggest that the socially optimal stringency of capital regulation must be identified using an economic risk-management logic: caring about inefficient systemic risk taking and its normally-invisible threat to macroeconomic stability. Standard macroeconomic variables (such as GDP or credit), evaluated at their unconditional means, or along paths in which the systemic shock does not realize, may give wrong indications about the desirable level of the capital requirements.[56]

6 Extensions and discussion

This section is structured in four parts. First, we analyze the transition from a low capital requirement regime to a higher target capital requirement regime, and assess the value of gradualism in approaching such target. Second, we consider state-dependent capital requirements and assess the effects of giving them a higher or lower pro- or counter-cyclical profile. Third, we discuss bailout policies, i.e. policies that attempt to reduce the contraction of credit supply after a systemic shock by transferring wealth to the bankers. Finally, we analyze the sensitivity of our quantitative results to changes in some of the parameters.

[56] In the results above, the optimal capital requirement seems to have a "large cost" relative to the suboptimal requirement when evaluated in terms of these variables.

6.1 Transitional dynamics and the value of gradualism

Prior sections have focused on economies in which the capital requirement γ_t remains constant over time, but the model can be extended to analyze the transition from a regime with some initial capital requirement γ_0 to a target requirement γ^* to be reached after T periods. To limit computational costs, we consider linear adjustment paths $\{\gamma_t\}$ with

$$\gamma_t = \gamma_0 + \frac{(\gamma^* - \gamma_0)}{T} t, \text{ for } t = 1, ...T - 1.$$

We set $\gamma_0 = 0.07$ so as to start from the low capital requirements regime of prior sections.

Figure 4 displays the social welfare (in permanent certainty-equivalent net consumption terms) associated with different target requirements γ^* when the economy starts in the pseudo-steady state induced by $\gamma_0 = 0.07$. Each curve corresponds to a value of γ^* ranging from 8% to 15% and the horizontal axis represents the number of transition periods T. The shortest transition, $T = 1$, implies announcing at $t = 0$ that the new requirement γ^* will come into effect at $t = 1$. With $T = 2$, the announcement means that half of the total increase in the requirement will take place at $t = 1$ and the remaining half at $t = 2$. And so on.[57]

Under our baseline calibration, the maximum welfare is obtained with a target capital requirement of 13% and 9 years of transition (although welfare is very similar with a target requirement of 12% and 5 years of transition).[58] Interestingly, if capital requirements were to be increased only up to a (suboptimal) level of 9%, then the net gains from reaching such target gradually rather than at once are virtually zero. For higher target requirements, however, maximizing (conditional) welfare requires transitional periods whose lengths increase with the target.

[57]To obtain the results, we first solve for equilibrium when the capital requirement is constant at the relevant target γ^* and use the obtained value functions, valid for $t \geq T$, together with backward induction on (5), to solve for the value functions relevant at each of the transitional periods $t = 1, 2, ...T - 1$. After obtaining the transitional value functions, we simulate 200 equilibrium paths of 1000 years each starting from the PSS of the economy with $\gamma = 0.07$, compute the welfare along each path, and report its average.

[58]Maximum welfare is not attained with a target equal to the unconditionally optimal requirement of 14% because of the assumed linearity of the adjustment path. This linearity implies an excessive cost to further increasing γ_t once it is sufficiently close to the unconditional optimum.

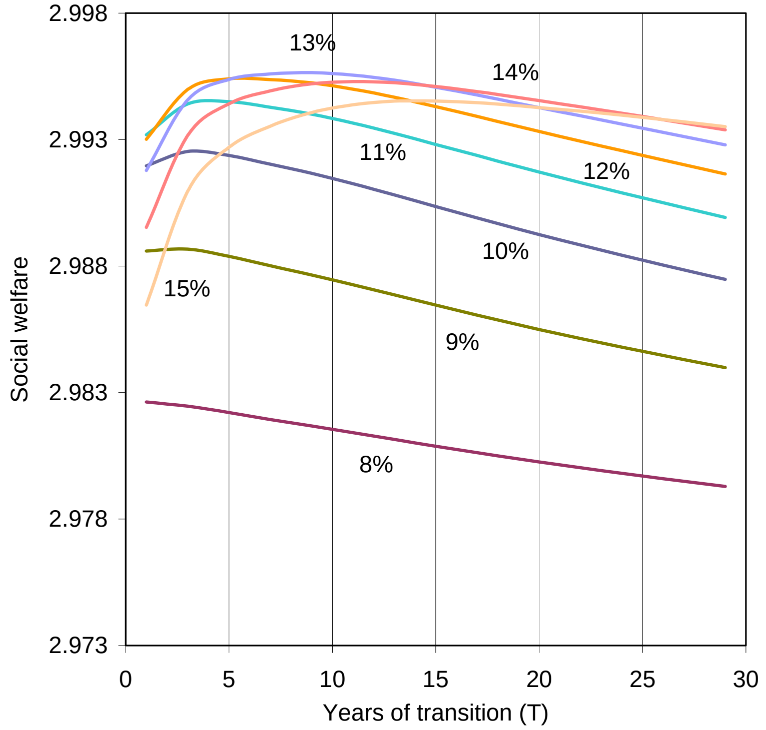

Figure 4 Social welfare for different target capital requirements as a function of the length of the transition (starting from the PSS with $\gamma = 0.07$)

Figure 4 also reconfirms the decreasing marginal social gains from increasing capital requirements in our economy. While there are significant gains from increasing capital requirements from 7% to up to 10% or 11% (even without gradualism), realizing significant gains from raising the target to, e.g., 12% (rather than 11%) requires gradualism (ideally 5 years). And the target of 13% only improves over the target of 12% if the transition is extended to more than 5 years (ideally 9 years).

6.2 Cyclically-adjusted capital requirements

We now turn to analyze the potential value of a cyclically-adjusted capital requirement, i.e. a potentially time-variant capital requirement $\gamma_t = g(e_t)$, where g is an arbitrary function taking values in the interval $[0, 1]$. Recent debates among bank regulators attribute virtuous counter-cyclical effects to making capital requirement tighter (looser) in good (bad) times, when equilibrium credit is likely to be above (below) what is socially desirable.[59] In our model this would imply having $g' > 0$ so that banks' lending capacity per unit of available bank capital e_t contracts when e_t is more abundant, and vice versa.[60]

We check the effects of cyclically-adjusted capital requirements by considering the following flexible functional form:

$$g(e_t) = \min\{\max[g_0 + g_1(\log(e_t) - \log(\overline{e})), 0], 1\}, \tag{17}$$

where g_0 and g_1 are constant parameters, and $\overline{e}$ is the amount of bank capital in the PSS of the economy with $\gamma_t = g_0$. According to this specification, the state-dependent capital requirement increases by about g_1 percentage points for each 1% difference between e_t and the reference value $\overline{e}$.

To find out which values of g_1 would make more sense under our baseline parameterization of the model, we set g_0 equal to the optimal unconditional time invariant requirement found in prior sections (14%) and compute the welfare attained under alternative values of g_1 around zero. Interestingly, we find that welfare increases smoothly as g_1 is set further and further *below* zero, down to a value of about $g_1 = -0.1$ (where the welfare gains relative to the time-invariant benchmark are of about 0.04%). In contrast, welfare falls quite sharply if g_1 is set further and further *above* zero (with a welfare loss of about 0.07% for $g_1 = 0.05$). This implies that the net ex ante welfare impact of relaxing capital requirements after a crisis is actually *negative* in our model. If anything, there would be ex ante social welfare gains from making capital effectively scarcer for low values of e.

The intuition for this striking result is that relaxing the capital requirement γ_t after a systemic shock reduces the last bank standing effect identified in prior sections and, hence, has a negative impact on bankers' incentives to preserve their capital after such shock. And it turns out that the welfare loss due to this higher ex ante systemic risk taking dominates

[59] Gersbach and Rochet (2012) and Malherbe (2013) identify setups where countercyclical bank capital requirements increase the efficiency of intertemporal investment decisions intermediated by banks.

[60] In parallel, it has been argued that making capital requirements tighter when bank capital is scarcer has undesirable procyclical effects (see, for instance, Repullo and Suarez, 2013).

the standard welfare gain from reducing the credit crunch once the systemic shock occurs. The last bank standing effect and the focus on systemic risk taking are unique features of our analysis, so this result constitutes a novel and important caveat on the policy prescriptions about the desirability of contercyclically-adjusted capital requirements obtained in other setups.[61]

6.3 Bailout policies

What does our model say about bailout policies? Are they at all desirable? Do they have undesirable side effects? In a spirit similar to the countercyclical capital requirements that we have just discussed, the rationale for bailout policies would be to try to avoid the sharp contraction in credit that follows the realization of a systemic shock. But the tool to expand credit supply after the shock would in this case be transferring wealth to bankers (rather than reducing the capital requirement).[62]

To structure the discussion that, in the interest of brevity we will perform in purely logical (rather than quantitative) terms, it is useful to distinguish between three classes of possible recipients of the bail-out transfers: (i) failed bankers, (ii) solvent bankers, and (iii) novel bankers. Among the first two classes, one may further distinguish between retiring and non-retiring bankers. In terms of the implications for the capital effectively available to banks after the bailout (and, hence, the resulting relief of the credit crunch), transfers to retiring bankers are clearly a waste, while all other transfers are perfect substitutes, since they contribute equally to reducing the ex post credit crunch.[63] But in addition to the direct effects on credit supply our analysis points to considering two important incentive effects.

First, the wealth transferred to the active bankers will reduce the equilibrium value of each unit of bank capital after the crisis. In terms of Figure 2, the $v(e)$ schedule will be-

[61] Having said that, the optimal policy suggested by our results faces a challenging time-consistency issue: while it is ex ante desirable to commit to keep capital requirements tight (or even tighter) after a shock, there may be positive ex post short-term welfare gains from relaxing the capital requirement immediately after the shock.

[62] Instrumenting the bailouts as wealth transfers to bankers is consistent with our maintained assumption that bank ownership requires possessing bank management talent, which is exclusive to bankers.

[63] These remarks on retiring bankers would correspond in practice to prescribing that banks receiving capital injections in a bailout should be subject to constraints on their payout policy.

come less steep, producing an increase in systemic risk taking (i.e. and upward shift in the $x(e)$ schedule). Second, there will also be direct incentive effects associated with bankers' prospects of receiving wealth transfers after the systemic shock (what media discussions about bank bailouts typically identify as "moral hazard"). From this last perspective, transfers to failed bankers would constitute a reward to systemic risk taking, those to solvent bankers would be a reward to systemic risk avoidance, and those to novel bankers would have no direct incentive effects.

Combining the three aspects of this discussion suggests that the socially preferable manner to implement a bailout would be to concentrate the wealth transfers on the active bankers with wealth invested in banks that resisted the systemic shock. At the cost of redistributing wealth from tax payers to bankers, it might well be the case that the relief of the credit crunch achieved in this manner, together with the direct incentive effects of the transfers to surviving bankers, offsets the negative indirect incentive effects coming from the reduction in the marginal value of bank capital after a systemic shock.

6.4 Sensitivity analysis

Figure 5 summarizes the comparative statics of the model relative to four of its key parameters. Each of its panels shows the effects of changing, one at a time, one of the parameters. We consider several alternative values around the baseline values reported in Table 1.

Each panel depicts three curves which describe the effects of the change in the parameter on the optimal time-invariant capital requirement (γ^*), the associated unconditional expected fraction of systemic loans, $E(x(e))$, and what that number would be if γ were held constant at the value of 14%. From left to right, from top to bottom, the four panels show the effects of moving the yearly unconditional probability (frequency) of the systemic shock (ε), the rate at which physical capital depreciates (or is destroyed) in failed firms (λ), the rate at which previously active bankers retire (ψ), and the (constant) size of the population of active bankers (ϕ).

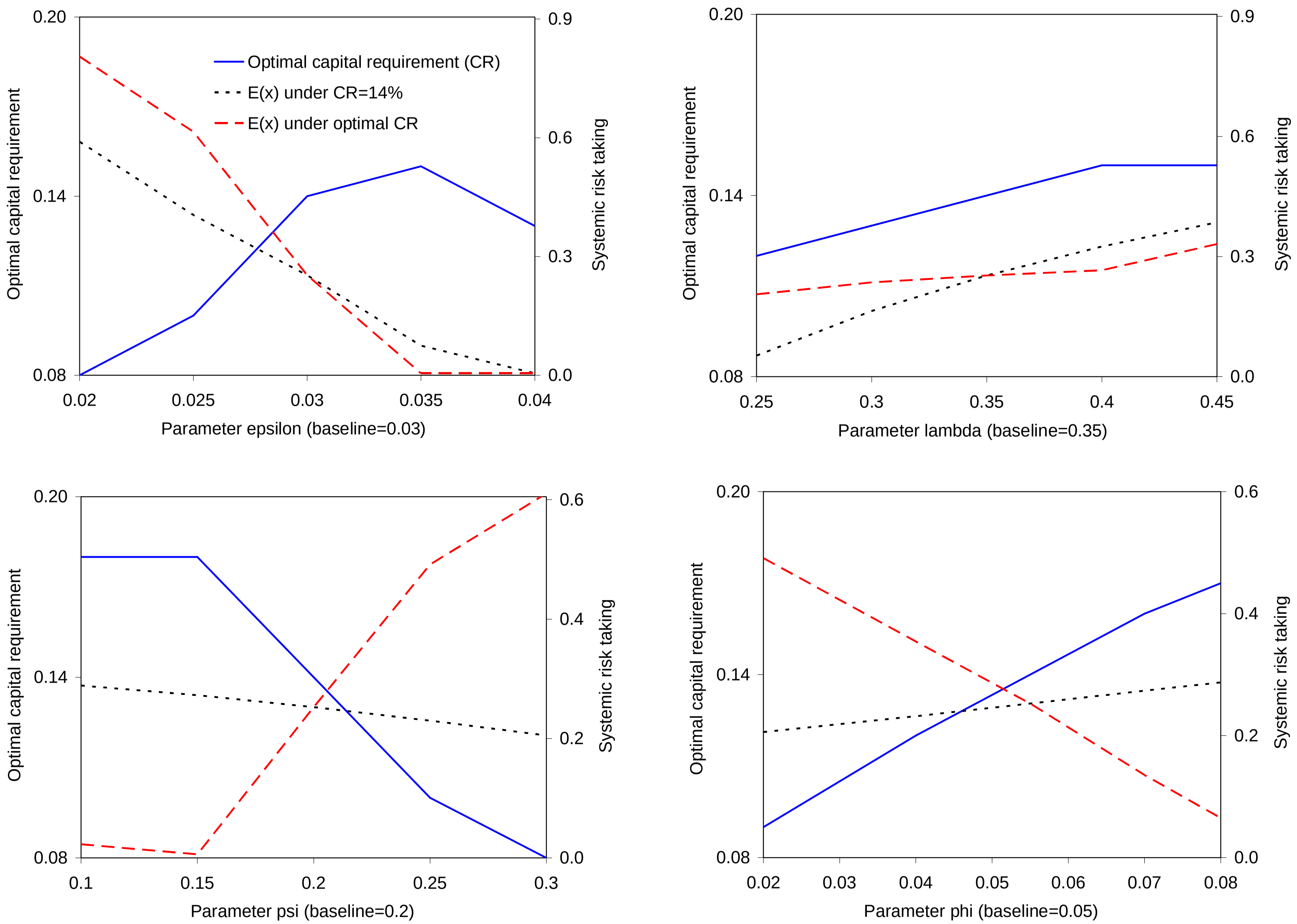

Figure 5 Dependence of the optimal γ (left axis) and systemic risk taking $E(x)$ (right axis) on some of the parameters

These results are generally self-explanatory so we will not discuss each of them in detail. However, it is worth commenting why the effect of the probability of suffering a systemic shock ε on the optimal capital requirement is non-monotonic. This is because, although systemic risk taking is a source of social losses that bankers do not fully internalize, the bankers themselves react to the larger risk of such shock by reducing their systemic risk taking. In fact, in the results reported in the figure, the optimal capital requirement increases in response to a rise ε up to the point in which systemic risk taking becomes zero. The non-monotonicity comes, quite intuitively, from the fact that, above that point, a larger ε reduces the γ^* required to keep bankers' systemic risk taking at a zero level.

7 Concluding remarks

This paper examines the effects of capital requirements on banks' endogenous systemic risk taking. We develop a dynamic general equilibrium model in which bankers decide on the (unobservable) exposure of their banks to infrequent systemic shocks. They trade off standard risk-shifting gains with the value of preserving their capital after these shocks. This value is enhanced by the existence of what we call a *last bank standing* effect: the high scarcity rents that surviving bank capital can earn after an important fraction of bank capital is wiped out by a shock.

We use the model to address some central issues in recent discussions on the macroprudential role of capital requirements. We find that capital requirements reduce credit and output in calm times but, by reinforcing the last bank standing effect, they also reduce systemic risk taking. The underlying trade-offs determine the existence of an interior social welfare maximizing level of the capital requirements. A second important implication of the last bank standing effect is that systemic risk taking increases as the economy expands, because aggregate bank capital becomes more abundant and, thus, earns lower rents. Another implication is that systemic risk taking may worsen if capital requirements are countercyclically adjusted, since such adjustment diminishes bankers' prospects of appropriating high scarcity rents by avoiding the exposure of their capital to the systemic shock.

Appendix

A Proof of Lemma 1

Bankers net payoffs from the portfolio of loans made by the non-systemic bank are, as described in problem (6), $(1-\pi_0)b_t + \pi_0(1-\lambda)k_t - (1+\rho)d_t$. Conditional on the information available at t, this payoff is deterministic since the default rate on non-systemic loans does not depend on the realization of the systemic shock (and idiosyncratic default risk is diversified away). Thus this whole expression can be taken out of the expectations operator in bankers' participation constraint, leaving it as follows:

$$E(v_{t+1})[(1-\pi_0)b_t + \pi_0(1-\lambda)k_t - (1+\rho)d_t] \geq q_t e_t. \tag{18}$$

Additionally, when investing in the non-systemic bank is optimal for bankers we have $q_t = E(v_{t+1}R_{0t+1}) = E(v_{t+1})R_{0t+1}$ since R_{0t+1} is not random. Hence (18) can be simplified to:

$$(1-\pi_0)b_t + \pi_0(1-\lambda)k_t - (1+\rho)d_t \geq R_{0t+1}e_t. \tag{19}$$

Under this writing, the required rate of return on equity at the non-systemic bank, R_{0t+1}, and the wage rate, w_t, are the sole channels through which the state of the economy at date t affects firm-bank decisions $(k_t, n_t, l_t, b_t, d_t, e_t)$.[64]

In the optimization problem stated in (6), both (19) and the constraint associated with the minimum capital requirement are binding.[65] So eventually the optimization problem involves six variables and four binding constraints. These constraints can be conveniently used to make substitutions that reduce the problem to one of unconstrained optimization with just two variables: k_t and n_t. Variables l_t, d_t, and e_t can be found recursively using the binding constraints $l_t = k_t + w_t n_t$, $d_t = (1-\gamma_t)(k_t + w_t n_t)$, and $e_t = \gamma_t(k_t + w_t n_t)$.

As for the loan repayment b_t, (19) and some further substitutions yield:

$$b_t = \frac{1}{1-\pi_0}\{[(1-\gamma_t)(1+\rho) + \gamma_t R_{0t+1}](k_t + w_t n_t) - \pi_0(1-\lambda)k_t\}. \tag{20}$$

Intuitively, the loan repayment at $t+1$ must compensate the bank in expected terms for the weighted average cost $(1-\gamma_t)(1+\rho) + \gamma_t R_{0t+1}$ of the funds $k_t + w_t n_t$ lent to the firm at

[64] Having a deterministic R_{0t+1} makes the system of equations that characterize equilibrium recursive by blocks. With a non-deterministic R_{0t+1} (say, if non-systemic firms also had some exposure to systemic shocks), the model would still be solvable but at a larger computational cost.

[65] To see the latter, notice that bankers can always invest their (generally scarce) wealth in insured deposits, implying $R_0 \geq 1+\rho$. With $R_0 > 1+\rho$ the constraint (19) is strictly easier to satisfy (so as to make the loans as cheap as possible) by funding any $l_t = d_t + e_t$ with the minimal possible e_t, which is $\gamma_t l_t$.

date t. The term $\pi_0(1-\lambda)k_t$ credits for the depreciated capital that the bank recovers when a firm fails.

Now, using (20) to substitute for b_t in the objective function of (6) gives rise to

$$\max_{(k_t,n_t)} (1-\pi_0)[AF(k_t,n_t)+(1-\delta)k_t]+\pi_0(1-\lambda)k_t-[(1-\gamma_t)(1+\rho)+\gamma_t R_{0t}](k_t+w_t n_t), \quad (21)$$

which is the reduced unconstrained maximization problem. The objective function of this problem is homogeneous of degree one in (k, n) so, like in neoclassical models with this feature, obtaining finite non-zero solutions requires the value of the objective function to be zero at the optimum. The first order conditions of the unconstrained problem (21) when evaluated at $n_t = 1$ (which is the aggregate 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.[66] Specifically, we obtain (8), which determines a k_t for each R_{0t+1}, and (9), which recursively determines a w_t for each R_{0t+1} and k_t.

The demand for bank capital that emanates from the above discussion is $e_t = \gamma_t l_t$ or, equivalently, $e_t = \gamma_t(k_t + w_t)$, as given by (10). Finally, we can obtain the expression for the loan rate that appears in (11) by using the definition $1 + r_t = b_t/l_t$, where b_t is found by evaluating (20) at $n_t = 1$ and $l_t = k_t + w_t$.■

B 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. Parameterize a possible state-dependent capital requirement as $\gamma_i = g(e_i)$, where $g(\cdot)$ is a given function. In the baseline calibrations, we have a constant requirement $\gamma_i = \gamma$ for all i, but in subsection 6.2 we use a more general function (see (17)).

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

[66] The 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 loans of the non-systemic bank.

conditions (8)-(10):

$$(1-\pi_0)[AF_k(k_i,1)+(1-\delta)]+\pi_0(1-\lambda)-[(1-\gamma_i)(1+\rho)+\gamma_i R_{0i}]=0, \tag{22}$$

$$(1-\pi_0)AF_n(k_i,1)-w_i[(1-\gamma_i)(1+\rho)+\gamma_i R_{0i}]=0, \tag{23}$$

$$\gamma_i(k_i+w_i)-e_i=0. \tag{24}$$

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

   (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-increasing 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^0_{1i} for each i :

$$(1-\pi_0)R^0_{1i}-(1-\pi_1)R_{0i}-\frac{1}{\gamma_i}(\pi_0-\pi_1)[(1-\gamma_i)(1+\rho)-(1-\lambda)\frac{k_i}{k_i+w_i}]=0. \tag{25}$$

8. Use the following extended version of (14) to find e^0_i and e^1_i for each i :

$$e^0_i=(1+\rho)\phi w_i+(1-\psi)\{[(1-x_i)R_{0i+1}+x_iR^0_{1i+1}]\widehat{e}_i+(1+\rho)(e_i-c_i-\widehat{e}_i)\}, \tag{26}$$

$$e^1_i=(1+\rho)\phi w_i+(1-\psi)[(1-x_i)R_{0i+1}\widehat{e}_i+(1+\rho)(e_i-c_i-\widehat{e}_i)]. \tag{27}$$

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

$$[(1-\varepsilon)v(e_i^0) + \varepsilon v(e_i^1)]R_{0i} - (1-\varepsilon)v(e_i^0)R_{1i}^0 \geq 0, \tag{28}$$

$$\{[(1-\varepsilon)v(e_i^0) + \varepsilon v(e_i^1)]R_{0i} - (1-\varepsilon)v(e_i^0)R_{1i}^0\}x_i = 0. \tag{29}$$

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

$$v(e_i) = \psi + (1-\psi)\beta[(1-\varepsilon)v(e_{i+1}^0) + \varepsilon v(e_{i+1}^1)]R_{0i+1}. \tag{30}$$

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.

C Social welfare

The patient agents who provide (insured) deposit funding to the banks at rate ρ break even in terms of their own net present value in all periods. Thus their net consumption flows make a zero net addition to social welfare W_t and we can safely leave them out in our welfare calculation. All other agents have a discount factor $\beta < 1/(1+\rho)$, which is the one at which we will 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). Focusing on the case in which bankers always reinvest all their accumulated wealth as bank capital, social welfare at any period t, W_t, can be expressed as

$$W_t = E_t\left(\sum_{s=0}^{\infty}\beta^s nc_{t+s}\right), \tag{31}$$

where

$$nc_t = -e_t + [1-(1+\psi)\phi]w_t + \beta\{y_{t+1} - (1+\rho)[d_t - (1+\psi)\phi w_t]\}, \tag{32}$$

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

$$gdp_{t+1} = [(1-x_t)(1-\pi_0) + x_t(1-u_{t+1})(1-\pi_1)]AF(k_t, 1), \tag{33}$$

$$\Delta_{t+1} = \delta + \{(1-x_t)\pi_0 + x_t[(1-u_{t+1})\pi_1 + u_{t+1}]\}(\lambda - \delta), \tag{34}$$

and $u_{t+1} \in \{0, 1\}$ indicates whether the systemic shock realizes ($u_{t+1} = 1$) or not ($u_{t+1} = 0$) at the end of period t. In this decomposition, nc_t is the present value of the net consumption flows that the impatient agents derive from the production period between dates t and $t+1$. Economic activity in that period initially absorbs bank capital e_t from the bankers and prepays wages w_t to all agents. However, some of these wages are not immediately consumed. Specifically, wages paid to the active bankers, ϕw_t, and to the workers who will be bankers at $t+1$, $\psi\phi w_t$, are saved 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 invested at date t, they bring about zero net consumption at date t.

At date $t+1$ (which explains the discount factor β in (32)), 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 $(1+\rho)[d_t - (1+\psi)\phi w_t]$ 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, (33) and (34), respectively, make clear that the GDP and the depreciation experienced by 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 u_{t+1}.

To further understand the sources of welfare in the expression above, it is useful to describe an alternative decomposition of the per-period welfare flow nc_t. This decomposition is based on the 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: $+[1-(1+\psi)\phi]w_t$.

2. Patient agents who act as depositors, who break even in NPV terms: $+0$.

3. Entrepreneurs, who in their role as producers break even state-by-state: $+0$.

4. Tax payers, who pay deposit insurance costs at $t+1$: $-\beta[(1+\rho)d_t-(1-\lambda)k_t]x_t u_{t+1}$.

5. Bankers who, as bank capital suppliers, contribute e_t at t and receive bank equity returns at $t+1$:
$$-e_t+\beta[(1-x_t)R_{0t+1}+x_t R_{1t+1}]e_t.$$

6. Active and next-period bankers who, as suppliers of labor at t, receive at $t+1$ the proceeds from having invested their wages in bank deposits: $+\beta(1+\rho)(1+\psi)\phi w_t$.

Notice that this decomposition is not explicit about bankers' net consumption. Along a full-reinvestment path bankers only consume when they cease in their activity. Their consumption flow is implicit in the two components referred to bankers (items 5 and 6). Specifically, the total income inflow assigned to old and new bankers at $t+1$ in the expressions above is $[(1-x_t)R_{0t+1}+x_tR_{1t+1}]e_t+(1+\rho)(1+\psi)\phi w_t$, while the only income outflow assigned to them at that date is the equity capital e_{t+1} contributed to banks for their next period of activity. Using (12), we obtain a net income flow of $\psi\{[(1-x_t)R_{0t+1}+x_tR_{1t+1}]e_t+(1+\rho)\phi w_t\}>0$ which indeed corresponds to the gross returns of the accumulated equity and the (deposited) last-period wages of the non-continuing bankers 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 (2011), "Systemic risk and macro-prudential regulation," New York University Stern School of Business, mimeo.

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 (2012), "Capital requirements and business cycles with credit market imperfections," *Journal of Macroeconomics* 34, 687–705.

Allen, Franklin, and Douglas Gale (2000), "Bubbles and crises," *Economic Journal* 110, 236–255.

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

Bernanke, Ben, and Mark Gertler (1989), "Agency cost, net worth and business fluctuations," *American Economic Review* 79, 14–31.

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.

Borio, Claudio (2003), "Towards a macroprudential framework for financial supervision and regulation?," *CESifo Economic Studies* 49, 181–215.

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

Brunnermeier, Markus K. and Yuliy Sannikov (2014), "A macroeconomic model with a financial sector," *American Economic Review* 104, 379–421.

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 (2013), "Risk shocks," NBER WP No. 18682.

Darracq Pariès, Matthieu, Christoffer Kok Sørensen, and Diego Rodriguez-Palenzuela (2011), "Macroeconomic propagation under different regulatory regimes: evidence from an estimated DSGE model for the Euro Area," *International Journal of Central Banking* 7, 49–113.

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

Diamond, Douglas, and Raghuram Rajan (2000), "A theory of bank capital," J*ournal of Finance* 55, 2431–2465.

Farhi, Emmanuel, and Jean Tirole (2012), "Collective moral hazard, maturity mismatch and systemic bailouts," *American Economic Review* 102, 60–93.

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

Gerali, Andrea, Stefano Neri, Luca Sessa, and Federico Signoretti (2010), "Credit and banking in a DSGE model of the Euro Area," *Journal of Money Credit and Banking* 42, 107–141.

Gersbach, Hans, and Jean-Charles Rochet (2012), "Aggregate investment externalities and macroprudential regulation," *Journal of Money, Credit and Banking* 44, 73–109.

Gertler, Mark, and Peter Karadi (2011), "A model of unconventional monetary policy," *Journal of Monetary Economics* 58, 17–34.

Gertler, Mark, and Nobuhiro Kiyotaki (2010), "Financial intermediation and credit policy in business cycle analysis" in Benjamin Friedman and Michael Woodford (eds.), *Handbook of Monetary Economics*, vol. 3, Elsevier, 547–599.

Gertler, Mark, and Nobuhiro Kiyotaki, and Albert Queralto (2012), "Financial crises, bank risk exposure and government financial policy," *Journal of Monetary Economics* 59, Supplement, S17–S34.

Gomes, Joao, Amir Yaron, and Lu Zhang (2003), "Asset prices and business cycles with costly external finance," *Review of Economic Dynamics* 6, 767–788.

Hanson, Samuel G., Anil K Kashyap, and Jeremy C. Stein (2011), "A macroprudential approach to financial regulation," *Journal of Economic Perspectives* 25, 3–28.

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

He, Zhiguo, and Arvind Krishnamurthy (2014), "A macroeconomic framework for quantifying systemic risk," NBER WP No. 19885.

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.

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

Iacoviello, Matteo (2005), "House prices, borrowing constraints, and monetary policy in the business cycle," *American Economic Review* 95, 739–764.

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.

Malherbe, Frederic (2013), "Optimal capital requirements over the business and financial cycles," London Business School, mimeo.

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

Meh, Cesaire, and Kevin Moran (2010), "The role of bank capital in the propagation of shocks," *Journal of Economic Dynamics and Control* 34, 555–576.

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 (2004), "Loan pricing under Basel capital requirements," *Journal of Financial Intermediation* 13, 496–521.

Repullo, Rafael, and Javier Suarez (2013), "The procyclical effects of bank capital regulation," *Review of Financial Studies* 26, 452–490.

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.