A Macroeconomic Model of Endogenous Systemic Risk Taking*

David Martinez-Miera
Universidad Carlos III

Javier Suarez
CEMFI and CEPR

First version: December 2010
This version: July 2012

Abstract

We analyze banks' systemic risk taking in a simple dynamic general equilibrium model. Banks collect funds from savers and make loans to firms. Banks are owned by risk-neutral bankers who provide the equity needed to comply with capital requirements. Bankers decide their (unobservable) exposure to systemic shocks by trading off risk-shifting gains with the value of preserving their capital after a systemic shock. Capital requirements reduce credit and output in "normal times," but also reduce banks' systemic risk taking and, hence, the losses caused by systemic shocks. Under our calibration of the model, optimal capital requirements are quite high, have a sizeable negative impact on GDP, do not require counter-cyclical adjustment, and should be gradually introduced.

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

JEL Classification: G21, G28, E44

* We would like to thank Matthieu Darracq Pariès, Giovanni Dell'Ariccia, Gianni De Nicolo, Martin Ellison, Carlos González-Aguado, Hans Gersbach, Frederic Malherbe, Claudio Michelacci, Enrico Perotti, Rafael Repullo, Diego Rodriguez-Palenzuela, Ruben Segura, Nicholas Trachter, Wolf Wagner, and seminar audiences at Banco de España, Bank of Canada, Bank of Japan, De Nederlandsche Bank, European Central Bank, HECER, International Monetary Fund, Kansas City Fed, New York Fed, Swiss National Bank, Universidad de Navarra, and the conferences on "Advances in Business Cycle Research – Directions Since the Crisis" in Brussels, "Financial Intermediation" in Funchal, "Financial Intermediation and the Real Economy" in Paris, "Macroprudential Policies, Regulatory Reform and Macroeconomic Modelling" in Rome, "Debt and Credit, Growth and Crises" in Madrid, FIRS 2012 in Minnesota, and RES 2012 in Cambridge 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).

Address for correspondence: CEMFI, Casado del Alisal 5, 28014 Madrid, Spain. Phone: +34-914290551. E-mail: suarez@cemfi.es.

1 Introduction

This paper analyzes the role of banks in generating endogenous systemic risk. We consider a canonical problem of excessive risk-taking by banks in a dynamic general equilibrium model. In our economy banks are subject to capital requirements and the supply of equity funding to them is limited by the wealth endogenously accumulated by bank owners. We find that systemic risk-taking, which can lead some banks to fail in equilibrium, is reduced by increasing capital requirements. Yet doing so has negative implications for the levels of credit and output, producing trade-offs relevant for determining the socially optimal level of the capital requirements.

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

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

[1] Systemic shocks resemble the *rare economic disasters* considered in Rietz (1988) and Barro (2009), among others. They may empirically correspond to phenomena such as the bust of the US housing market (and its implications for subprime mortgages, securitization markets, and money markets) around the Summer of 2007. We focus on banks' endogenous exposure to *exogenous* systemic shocks to avoid the complexity associated with modeling the mechanisms that generate correlated losses (bubbles, negative spillovers caused by fire sales, interbank linkages, bank panics, etc.). A recent attempt in that direction is Brunnermeier and Sannikov (2011).

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

[3] See Section 2 for a review of the literature more closely related to our model.

independent failures. In contrast, systemic production processes fail simultaneously if the negative systemic shock occurs (and in an independent fashion if it does not).

Following the literature on risk shifting, we assume that systemic firms have probabilities of failure that, conditional on not suffering the systemic shock, are lower than those of non-systemic firms. However, unconditionally systemic firms are more likely to fail and, thus, generate lower expected net present value. Yet they may be a tempting investment opportunity for highly levered institutions such as banks, which enjoy limited liability and/or have the ability to issue liabilities protected by safety net guarantees.

Following many models in the literature on bank risk taking and capturing features observed in reality, we assume that (i) bank deposits are fully insured by the government and (ii) banks are subject to capital requirements. We model the dynamics of bank capital along the same lines as Gertler and Kiyotaki (2010), Gertler and Karadi (2011), and Meh and Moran (2010). This implies thinking of bank capital as the funds provided by the special class of agents who own and manage the bank (the "bankers"). The maximun available amount of this funding is determined by bankers' wealth, whose endogenous dynamics is affected by the profits and losses made by the banks in each period.[4]

Bankers may be interested in making systemic loans because they are protected by limited liability and, if their bank fails, part of the losses go to the government as the provider of deposit insurance.[5] Bank capital requirements influence bankers' incentives in regards to the adoption of systemic risk in two ways. First, the conventional leverage-reduction effect diminishes the static gains from risk-taking. Second, higher capital requirements increase the relative scarcity of bank capital in each state of the economy, altering bankers' dynamic incentives in a interesting manner.

Indeed, the anticipated value of a unit of bank capital in the future is key to systemic risk-taking decisions. When a banker's capital is devoted to make systemic loans, the banker obtains higher returns insofar as the systemic shock does not occur. But if the systemic shock

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

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

occurs, the invested equity is lost. In contrast, if the banker invests in non-systemic loans, he receives a lower return in "normal times" (i.e. if the systemic shock does not realize), but preserves his capital when the systemic shock occurs. The destruction of bank capital after a shock allows surviving capital to earn higher scarcity rents and produces a "last bank standing effect" like in Perotti and Suarez (2002). This effect reduces bankers' inclination for systemic lending.[6]

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

Strengthening capital requirements reduces bankers' systemic risk taking through the two channels mentioned above. Higher capital requirements reduce the proportion of resources going into inefficient systemic investments, which in turn reduces the loss of bank capital and the contraction in credit supply produced by the possible realization of the systemic shock. However, these gains come at the cost of reducing credit and output in "normal times." Measuring welfare as the expected present value of total net consumption flows in the economy, we find that there is a unique interior social welfare maximizing level of capital requirements.

Under our calibration, social welfare is maximized under a positive and relatively large capital requirement: it is optimal to require banks to finance 14% of their loans with bankers' wealth. To fix ideas, we compare the scenario with a 14% capital requirement with another with a 7% capital requirement. We find that the unconditional mean of the fraction of bank equity devoted to support systemic lending is 71% under the low capital requirement

[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; partial dilution would lower the effectiveness of this mechanism. Full dilution is, in principle, compatible with resolution practices in which the failed banks continue as a going concern (but in hands of new owners). In our model, however, banks have no going-concern value beyond the value of the equity of their owners.

and 24% under the optimal capital requirement. The social welfare gain from having the requirement of 14% rather than 7% is equivalent to a perpetual increase of 0.9% in aggregate net consumption.

Importantly, common macroeconomic aggregates such as GDP and bank credit have lower unconditional expected values (6.5% and 21% lower, respectively) in the economy with a capital requirement of 14% than with one of 7%. However, the fall in aggregate net consumption, GDP, and bank credit in the year that follows a systemic shock is much lower with the 14% capital requirement.[7]

The model is suitable for the analysis of the transition from a regime with a low capital requirement (say, 7%) to another with a higher capital requirement. It allows to explicitly take into account transitional dynamics and 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. We find that it is socially optimal to implement the higher requirements in a gradual way (over 7 to 10 years under our calibration) and to establish a more modest long-term goal than if transitional cost were neglected (12% or 13%, depending on the desired speed of convergence to the target, rather than 14%).

The rest of the paper is organized as follows. Section 2 places the contribution of the paper in the context of the existing literature. Section 3 describes the model. Section 4 derives the conditions relevant for the definition of equilibrium. Section 5 describes our calibration exercise and the main quantitative results. Section 6 presents the results on optimal 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.

[7] In the presence of systemic risk, GDP is a bad proxy of social welfare since it neglects important losses associated with the realization of such risk. Specifically, the value added measured in GDP is gross of capital depreciation and, hence, does not properly account for the depreciation of physical capital associated with firm failure (which occurs with a larger unconditional probability under the systemic production mode).

2 Related literature

Our paper is related to recent efforts to incorporate banks and their (endogenous) contribution to systemic risk into core macroeconomic analysis. Dynamic stochastic general equilibrium (DSGE) models in use by central banks prior to the beginning of the crisis (e.g. in the tradition of Smets and Wouters, 2007) paid no or very limited attention to financial frictions. Several models considered idiosyncratic default risk and endogenous credit spreads using the framework provided by Bernanke et al. (1999) but very few were explicit about banks.[8] More recently, various authors have extended models in the DSGE tradition with the explicit goal of capturing banking frictions.[9] However, the reduced-form approach typically leaves aside an explicitly microfounded role for the introduced regulatory ingredients and impedes a fully-fledged welfare analysis.

The papers more closely related to our modeling of bank capital dynamics are Gertler and Kiyotaki (2010), Gertler and Karadi (2011), and Meh and Moran (2010), which also postulate an explicit (albeit different) connection between bank capital and bankers' incentives. These papers prescribe for bankers' wealth the same type of dynamics as for entrepreneurial net worth in the models such as those of Carlstrom and Fuerst (1997) and Kiyotaki and Moore (1997). In Gertler and Kiyotaki (2010) and Gertler and Karadi (2011), like in Hart and Moore (1994), bankers have to fund some minimal fraction of their banks with their own funds 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. in a setup in which banks' outside financiers are not protected by government guarantees and bankers make costly unobservable decisions regarding the monitoring of their borrowers.

Similar bank capital dynamics (and rationale for capital requirements) are postulated by Brunnermeier and Sannikov (2011), who put the emphasis on identifying channels through

[8] An exception is Van den Heuvel (2008), who introduces banks' liquidity provision function in a macroeconomic setup and assesses the welfare cost of capital requirements. Yet, in his model, capital requirements are an ad hoc piece of regulation with no explicit impact on risk-taking incentives or bank solvency.

[9] See, for instance, Agénor et al. (2009), Christiano et al. (2010), Darracq Pariès et al. (2011), and Gerali et al. (2010).

which a sequence of small shocks can lead to a crisis. The paper captures a rich interaction between value-at-risk based capital requirements, fire sales, and asset price volatility but does not discuss optimal capital regulation.[10] Like us and differently from most other papers, the authors consider the full stochastic non-linear dynamics of the model rather than a linear approximation to some non-stochastic steady state.

Our explicit focus on banks' risk taking decisions, and on how regulatory capital requirements interfere with them, 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 at al. (1998) and Freixas and Rochet (2008) provide excellent surveys of subsequent contributions.

Excessive risk-taking by banks 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.[11] The role of capital requirements in ameliorating this problem and their interaction with the incentives coming from banks' franchise values is a central theme in Hellmman at al. (2000) and Repullo (2004), where banks earn rents due to market power.[12] 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 (2007, 2008). The shadow value of bank capital in our context plays an incentive role similar to that of franchise value in the previous literature. However, differently from the prior tradition, the banks in our model are perfectly competitive and the relevant

[10]The analysis focuses on the polar case in which capital requirements guarantee that banks never fail.

[11]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.

[12]We abstract, for simplicity, from the entrepreneurial-incentives channel emphasized by Boyd and De Nicoló (2005) in their reexamination of the link between market power and bank solvency (see also Martinez-Miera and Repullo, 2010). Extending our framework for the explicit consideration of entrepreneurs' incentives (and the dynamics of their net worth) would require a two-tier formulation similar to Holmström and Tirole (1997) which, in its simpler formulation, would add a second aggregate state variable in the model, making its solution computationally more demanding.

continuation value is attached to bank capital, which is solely provided by bankers and 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 one period and is subject to failure risk. Banks intermediate between savers and firms so as to allow the latter to pay for their factors of production in advance, and banks are owned by bankers who provide them with the equity needed to satisfy 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 r per period, and can be interpreted as the exogenous return on some risk-free technology. Patient 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.[13]

Impatient agents, of whom there is a continuum of measure one, are infinitely lived, have a discount factor $\beta < 1/(1+r)$, and inelastically supply a unit of labor per period at the prevailing wage rate w_t. Most impatient agents are mere *workers*. Each worker has a small independent probability $\phi\psi/(1-\phi)$ of learning in each date t that he will become a *banker* (i.e. posses the skills needed to own and manage a bank) at date $t+1$. In parallel, each banker active at date t has a small independent probability ψ of becoming a mere worker again at date $t+1$. This produces a stationary size ϕ for the population of active bankers.

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

Finally, a tiny fraction μ of the impatient agents who do not act as bankers in each given date receive the opportunity to act as *entrepreneurs* (i.e. owning and managing a firm) during the imminent period. We focus on parameterizations under which impatient agents find it optimal to act as bankers or entrepreneurs if the occasion arises.[14] We also assume that the probabilities $\phi\psi$ and μ are small enough for the accumulation of wealth by workers not to be worthy prior to learning about their conversion into bankers or entrepreneurs.[15]

3.2 Firms

The entrepreneurs active in every period run a continuum of perfectly competitive firms indexed by $i \in [0, \mu]$. 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 $t + 1$.[16] 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.[17] 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 $\alpha \in (0, 1)$.[18] In case of failure, firms do not produce any output on top of depreciated capital.

[14] In equilibrium, entrepreneurs eventually receive a competitive profit of zero at all dates.

[15] Such wealth accumulation will expand the number of state variables in the model, complicating the quantitative analysis.

[16] 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).

[17] 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.

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

The possible correlation of z_{it+1} across firms is due to the exposure of firms to a common *systemic* shock $\varepsilon_{t+1} \in \{0,1\}$, whose bad realization $\varepsilon_{t+1} = 1$ is assumed to occur with a constant independent small probability ε at the end of each period.[19] The production technology can be operated in two modes that differ in their degree of exposure to the systemic shock: one is not exposed to it or *non-systemic* ($\xi_{it} = 0$), while the other is totally exposed to it or *systemic* ($\xi_{it} = 1$).

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

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

so, by the law of large numbers, the failure rate associated to any positive measure of non-systemic firms is constant and equal to p_0.

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

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

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

$$z_{t+1} = \begin{cases} p_1 & \text{if } \varepsilon_{t+1} = 0, \\ 1 & \text{if } \varepsilon_{t+1} = 1, \end{cases} \tag{3}$$

since systemic firms fail independently (with probability p_1) if the negative systemic shock does not occur and simultaneously if it occurs.

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

A1. $E(z_{it+1} \mid \xi_{it} = 1) = (1-\varepsilon)p_1 + \varepsilon > E(z_{it+1} \mid \xi_{it} = 0) = p_0.$

A2. $p_0 > p_1.$

[19] The use of ε to describe the probability with which the dichotomous systemic shock realizes (i.e. $\varepsilon_{t+1} = 1$) helps simplify some notation below. Specifically, we will use the superscripts ε and $1-\varepsilon$ to identify the realizations of random variables contingent, respectively, on the realization ($\varepsilon_{t+1} = 1$) or not ($\varepsilon_{t+1} = 0$) of the systemic shock.

Assumption A1 means that systemic firms are overall less efficient (i.e. yield lower total expected returns) than non-systemic ones, so systemic risk taking is socially undesirable. However, assumption A2 implies that lending to systemic firms may be attractive to bankers protected by limited liability, who would enjoy less defaults insofar as the systemic shock does not realize and suffer losses limited to their initial capital contributions otherwise.[20]

The entrepreneurs, who run the firms, are penniless, enjoy limited liability, and maximize their expected payoffs at the end of the production period, when they become mere workers again.[21] 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 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.[22] The tuple $(\xi_{it}, k_{it}, n_{it}, l_{it}, B_{it})$ is determined in the contracting between each firm and its bank at date t, where, reflecting bank competition, entrepreneurs have all the bargaining power.[23] Importantly, a firm's systemic orientation ξ_{it} is private information of the firm and its bank, ruling out regulations directly contingent on it.

3.3 Banks

Regulation obliges banks to finance a fraction γ_t of their one-period loans to firms with equity capital i.e. funds coming from bankers' accumulated wealth. Banks complement their funding with fully-insured one-period deposits taken from patient agents (and the would-be bankers who save their labor income so as to invest in bank capital in the next date).[24] The deposit insurance scheme is paid for with contemporaneous non-distortionary taxes levied

[20] It can be shown that with $p_1 \geq p_0$ no bank would ever get involved in the funding of systemic firms.

[21] 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.

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

[23] Nevertheless, as discussed in Section 4.2, given the constant returns-to-scale technology and the competitive product and factor markets, entrepreneurs' equilibrium profits will end up being zero in all states.

[24] We assume that impatient agents cannot borrow for 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 the patient agents and firms from the banks because they own assets at the end of each period (depreciated physical capital and net output) that are *pledgeable*.

on impatient agents.[25]

We assume that banks hold *non-granular* loan portfolios, that is, extend infinitesimal loans to a continuum of firms, thus fully diversifying away firms' idiosyncratic failure risk.[26] Diversification, however, does not eliminate the systemic risk associated with the 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.[27] 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* ($\xi = 0$) and a representative *systemic bank* ($\xi = 1$). Each bank's balance sheet constraint imposes

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

for $\xi = 0, 1$, where $l_{\xi t}$ denotes the loans made by the bank at date t, $d_{\xi t}$ are its deposits, and $e_{\xi t}$ is equity provided by bankers.[28]

The allocation of bank capital to each bank takes place in a perfectly competitive fashion. Bankers can invest their wealth at any date t as capital of the non-systemic bank, capital of the systemic bank or insured deposits, and they can also consume all or part of their wealth.[29] If they contribute $e_{\xi t}$ to bank ξ, 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 expectation about bank equity returns (and the value of the resulting wealth) across different possible states at $t+1$.

Banks take as given bankers' valuation of wealth across possible states at $t+1$ and their *value-weighted required return on wealth* (see section 4.1). Based on this, they formulate

[25]E.g. a tax on 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. Deposit insurance simplifies the analysis of risk-shifting but is not essential for our results insofar as the pricing of uninsured deposits cannot be contingent on the unobservable systemic orientation of each bank.

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

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

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

[29]Bankers can opt for any mixture of these four options so 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.

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

4 Equilibrium analysis

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

In order to facilitate the exposition, we will focus the presentation of equilibrium conditions in the main text on the case in which bankers invest their whole wealth as equity of the existing banks (*full reinvestment* equilibrium). In Appendix C, we generalize these conditions to cover equilibria in which bankers consume part of their wealth or save part of it in the form of bank deposits.

We will assume that banks play a pooling equilibrium in which the non-systemic bank solves its individual maximization problem when contracting with firms, while the systemic bank prevents being identified as such (which would imply its dissolution by the regulator) by mimicking the non-systemic bank in every aspect except the unobservable systemic orientation of the firms receiving its loans ($\xi_{it} = 1$). Firms, in turn, will be indifferent in equilibrium between adopting a systemic or non-systemic orientation because the presence of competitive factor and product markets, together with their constant returns to scale technology, will make their equilibrium profits equal to zero.

Importantly, when the systemic shock does not occur, the realized return on equity will tend to be higher at the systemic bank (R_{1t+1}) than at the non-systemic bank (R_{0t+1}), but we assume that bank accounts and managerial compensation practices are opaque enough to allow bankers to appropriate the excess return without being discovered.[30] 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.[31]

4.1 Bankers' portfolio problem

Continuing bankers (i.e. bankers active at date t who do not convert back into workers at date $t+1$) 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.[32]

When a banker converts into a worker, which happens with probability ψ, his only alternatives are either to save the wealth as a bank deposit (earning a gross return $1+r$ 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 will assume that consuming is the optimal decision and, thus, the value of one unit of his wealth is just 1.[33]

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

[30] 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. Another observable characteristic that might help differentiate the systemic bank from the non-systemic bank, if there were just one bank of each type, would be size. But the underlying constant returns to scale technologies are compatible with having many banks within each class and arbitrary sizes for each bank.

[31] This is consistent with the view in Diamond and Rajan (2000).

[32] 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_te_t.

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

Bellman equation for v_t:

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

The terms within the first max operator reflect, in this order, the following possibilities: (i) consuming the wealth, (ii) investing in deposits, (iii) investing in equity of the non-systemic bank (which yields a gross return R_{0t+1}), and (iv) investing in equity of the systemic bank (which yields a gross return R_{1t+1}).

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 (rather than consuming it) if $(1+r)E_t(v_{t+1}) \geq 1$ and the returns on bank equity (R_{0t+1} or R_{1t+1}) are small enough, i.e. $(1+r)E_t(v_{t+1}) \geq \max\{E_t(v_{t+1}R_{0t+1}), E_t(v_{t+1}R_{1t+1})\}$. 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).[34]

For brevity, the equilibrium conditions presented in the main text focus on the case with $\beta\max\{E_t(v_{t+1}R_{0t+1}), E_t(v_{t+1}R_{1t+1})\} > \max\{1, \beta(1+r)E_t(v_{t+1})\}$. Then active bankers' optimal portfolio decisions can be described as follows:

- Invest all wealth in equity of the non-systemic bank if $E_t(v_{t+1}R_{0t+1}) > E_t(v_{t+1}R_{1t+1})$.
- Invest all wealth in equity of the systemic bank if $E_t(v_{t+1}R_{1t+1}) > E_t(v_{t+1}R_{0t+1})$.
- Invest in equity of any of the banks 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. To avoid problems interpreting the pooling equilibrium in which the systemic bank mimics the non-systemic bank in all dimensions (except in setting $\xi_{it} = 1$ for all its funded firms), we will focus on parameterizations under which the equity of the

[34] The Cobb-Douglas production technology and the Walrasian determination of equilibrium wages tends to make the supply of loans infinitely profitable when the amount of supplied loans tends to zero.

non-systemic bank is always sufficiently attractive to bankers in equilibrium, in which case $Q_t = E_t(v_{t+1}R_{0t+1})$ for all t.[35]

4.2 Firm-bank lending contracts

This subsection describes how the non-systemic bank sets the terms of the lending relationship with each of its funded firms. The systemic bank will just mimic all these terms, except the unobservable systemic orientation of its funded firms. The contract signed between the non-systemic bank and each of its funded entrepreneurs will set $(\xi_{it}, k_{it}, n_{it}, l_{it}, B_{it}) = (0, k_t, n_t, l_t, B_t)$, where k_t, n_t, l_t, and B_t solve the following problem:[36]

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

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. 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 his firm does not fail, and zero when it fails.

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 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, $(1-p_0)B_t$, plus the payment coming from the recovery of depreciated physical capital in failed firms, $p_0(1-\lambda)k_t$, minus the payments due to depositors, $(1+r)d_t$.

[35] It is possible to analytically show that having a small measure of active bankers ($\phi \to 0$) or low risk-shifting incentives ($p_1 \to (p_0-\varepsilon)/(1-\varepsilon)$) is sufficient to rule out equilibria with all bankers' wealth is invested in equity of the systemic bank ($x_t = 1$). Intuitively, with no entry of new bankers ($\phi = 0$), 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 the non-systemic manner.

[36] As in neoclassical production theory, the constant returns-to-scale technology will make the optimal size of individual firms (and, hence, of individual loans) undetermined in equilibrium. So 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.

The last three constraints in problem (6) reflect (i) the use of loans to pay firms' capital and labor in advance, (ii) the balance sheet identity $l_t = d_t + e_t$, and (iii) the regulatory capital requirement $e_t \geq \gamma_t l_t$.

The fact that equity returns at the non-systemic bank are deterministic allows us to divide both sides of the first constraint in (6) by $E(v_{t+1})$ and obtain

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

where R_{0t+1} is to be thought of the market-determined "required" return on equity at the non-systemic bank (taken as given by banks). For $R_{0t+1} > 0$, this participation constraint implies that the bankers' (deterministic) net payoffs from investing in the non-systemic bank are always positive in equilibrium.

Importantly, 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. If it were strictly positive, entrepreneurs would like to scale their firms up to infinity; if it were strictly negative, they would not find it feasible to operate their firms at positive scale.

Expressing the 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 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}, optimal firm-bank lending contracts and labor market clearing imply that, in a pooling equilibrium:*
(a) firms' aggregate demand for physical capital k_t satisfies

$$(1 - p_0)[AF_k(k_t, 1) + (1 - \delta)] + p_0(1 - \lambda) = (1 - \gamma_t)(1 + r) + \gamma_t R_{0t+1}, \tag{8}$$

(b) the market clearing wage rate w_t satisfies

$$(1 - p_0)AF_n(k_t, 1) = [(1 - \gamma_t)(1 + r) + \gamma_t R_{0t+1}]w_t, \tag{9}$$

(c) the aggregate demand for equity capital e_t satisfies

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

(d) the loan rate r_{Lt} satisfies

$$1 + r_{Lt} = \frac{1}{1-p_0}\{[(1-\gamma_t)(1+r) + \gamma_t R_{0t+1}] - p_0(1-\lambda)\frac{k_t}{k_t + w_t}\}. \tag{11}$$

Equations (8) and (9) reflect how the production problem solved by banks and firms in our economy extends the canonical problem of perfectly-competitive firms in static production theory. 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 funding bank's weighted average cost of funds, $(1-\gamma_t)(1+r) + \gamma_t R_0$.

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.[37] 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 all dates.

4.3 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 capitalized value $\phi(1+r)w_t$ of the labor income earned by currently active bankers in the prior date (which they are assumed to save), and (ii) the gross returns on the wealth

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

$(1-\psi)e_t$ that continuing bankers invested as bank capital at date t.[38] This results in the following law of motion for e_{t+1}:

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

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 ε_{t+1}. From A2, if the systemic shock does not realize, one unit of capital of the systemic bank yields the gross return $R_{1t+1}^{1-\varepsilon} > R_{0t+1}$, where

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

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 p_1 rather than p_0.

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 $R_{1t+1}^{\varepsilon} = 0 < R_{0t+1}$.[39]

From date t perspective, the aggregate bank capital available at date $t+1$ is a dichotomous random variable whose law of motion can be expressed as:

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

which clearly shows its dependence of the aggregate shock ε_{t+1}.

Looking back at (5) and using (14), it is immediate to summarize the conditions for the compatibility of particular values of x_t with bankers' optimal portfolio decisions.

[38] Appendix C states equilibrium conditions for the general case in which active bankers may find it optimal to consume part of their wealth or to keep part of it inverted as bank deposits. For simplicity, the assumption that bankers save all their labor income will not be relaxed, but we will check that such behavior is always optimal under our parameterization of the model.

[39] 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 *In equilibria with $x_t \in [0,1)$ we must have:*

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

Moreover, if (15) holds with equality, then any $x_t \in [0,1)$ is compatible with bankers' optimization, otherwise only the corner solution $x_t = 0$ is compatible.

The corner solution without systemic risk-taking ($x_t = 0$) that emerges when (15) holds with strict inequality can be formally captured by imposing:

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

which can be interpreted as a complementary slackness condition.

4.4 Equilibrium

Along a full-reinvestment equilibrium, the state of the economy at any date t (right before investment decisions are made for one more period) can be summarized by a single state variable: the total wealth available to the active bankers e_t. The stochastic evolution of e_t is driven by the realization or not of the systemic shock at the end of every period as described in (14).

The equilibrium values of other variables in the model can be thought of as functions of the state variable e_t. These functions must satisfy the individual optimization and market clearing conditions 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 $(v(e), x(e), k(e), w(e), R_0(e), R_1^{1-\varepsilon}(e))$ describing the key endogenous variables as functions of $e \in [\underline{e}, \overline{e}]$, such that all the sequences $\{e_t\}_{t=0,1,...}$ and $\{v_t, x_t, k_t, w_t, R_{0t+1}, R_{1t+1}^{1-\varepsilon}\}_{t=0,1,...}$ that they generate satisfy:*

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

2. *The clearing of all markets.*

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

Thus, 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 $R_{1t+1}^{1-\varepsilon}$ that arise when aggregate bank capital is $e_t = e$ will be found by evaluating the various components of the tuple $(v(e), x(e), k(e), w(e),\, R_0(e), R_1^{1-\varepsilon}(e))$.

Appendix C relaxes requirement 3 in the definition of equilibrium, so as to allow for solutions in which bankers find it optimal to consume part of their wealth or to invest part of it in bank deposits in some states. Appendix C also describes the numerical solution method that we use to compute the equilibrium in the quantitative part.

4.5 Key equilibrium forces

This subsection is going to intuitively explain some of the key mechanisms at work in the determination of equilibrium. Given the fixed supply of labor and the underlying constant-returns-to-scale technology, the aggregate returns to bank lending are marginally decreasing. This makes one unit of bankers' wealth (which can be used to expand banks' lending capacity) more valuable when bankers' aggregate wealth is more scarce. Specifically, we have $v'(e) < 0$ if bankers invest all their wealth in bank equity and $v'(e) = 0$ if, at some point, they devote it, in the margin, to alternative uses.

Intuitively, when bankers find it optimal to invest all their wealth in bank equity, increasing e expands banks' lending capacity, loans become cheaper, and firms expand their activity, which in equilibrium, after wages adjust, implies devoting more physical capital to production. Like in the neoclassical growth model, the fixed supply of labor makes the aggregate return on physical capital marginally decreasing. This implies that the marginal value of bank lending and, hence, the scarcity rents appropriated by bank capital are decreasing in e.[40]

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

In combination with the dynamics of bank capital described in (14), the aggregate marginally decreasing returns to e imply that, after sufficiently many periods without suffering a systemic shock, the economy will converge 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 not realize (but agents are fully aware of the possible occurrence of such shock at any date). If the shock realizes, all the e invested as equity of the systemic bank is lost and the process of accumulation of bankers' wealth and the convergence to the PSS (possibly disrupted by the arrival of another systemic shock) starts over again.

To understand the intuition driving bankers' systemic risk-taking decisions, represented by x_t, the crucial equations are (15) and (16). As previously mentioned, (13) implies $R_{1t+1}^{1-\varepsilon} > R_{0t+1}$ so satisfying (15) requires a sufficiently large value of $\varepsilon v(e_{t+1}^{\varepsilon})$, i.e. a high valuation for the equity that survives the systemic shock. Since $v(e)$ is decreasing, this in turn requires a sufficiently low value of e_{t+1}^{ε}. Intuitively, the bankers who give up the gains from risk-shifting must be compensated by the expectation of obtaining a large revaluation of their wealth when their bank survives the systemic shock.

By (14), a larger x_t implies, other things equal, a larger aggregate loss of bank capital when the shock occurs, and hence a lower e_{t+1}^{ε} and a larger $v(e_{t+1}^{\varepsilon})$. This establishes a self-equilibrating mechanism for the determination of x_t. Equation (16) embeds the indifference (or no-arbitrage) condition required for producing an interior $x_t \in (0, 1)$.

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 and decomposed in various forms, depending on the dimension along which the relevant overall aggregation is performed. One can infer the net consumption flow that the economy generates for each class of agents in each date t and aggregate across agents. Alternatively, one can just look at the differences between the aggregate quantity of the consumption good the economy produces at the end of a period and the quantity which is reutilized as a factor of production (physical capital) in the next period. Appendix B provides an explicit expression for W_t (equation (22)) and an associated flow measure of welfare ω_t, which are explained there using two intuitive decompositions.

Following convention, we will describe below the gains and losses in expected welfare, $E(W_t)$, as percentage differences in the certainty-equivalent *permanent* aggregate net consumption *flow* that would give rise to such welfare, which can be calculated as $(1-\beta)E(W_t)$.[41]

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) with those obtained with the requirement of 14% (γ=0.14) that, under the parameterization presented in Table 1, maximizes the *unconditional* expected value of W_t.[42]

In Section 6, we analyze some cases of 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 Calibration

Our quantitative results are based on assuming that one period in the model corresponds to one year in calendar time. Table 1 contains the parameters chosen for our calibration of the model. We have tried numerous other parameter configurations to check the performance of the solution method and the generality of the qualitative results, in both cases with positive results. Some of these robustness checks will be discussed in Section 6.

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

[42] "Unconditional" means that the starting points of the simulated paths over which we compute $E(W_t)$ are extracted from the ergodic distribution of e_t in proportion to the relevant probabilities. The support of the ergodic distribution is made of the values reached with strictly positive probability by e_t along sufficiently long histories of the economy. Our results are based on simulating one path of 10,000 periods.

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

Patient agents' discount rate	r	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	p_0	0.03
Idiosyncratic default rate of systemic firms	p_1	0.018
Probability of a systemic shock	ε	0.03
Bankers' exit rate	ψ	0.20
Fraction of wages devoted to forming new bank capital	ϕ	0.05

The model is 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), whose probability of occurring ε is one of the parameters. The discount rate of the patient agents r is chosen equal to 2% to capture a situation with low real interest rates such as the one observed in developed economies in the years leading to the 2007 financial crisis. The discount factor of the impatient agents embeds a discount rate which is approximately twice as large as r. In the literature on external financing frictions it is standard to assign values of this order of magnitude, or even larger, to their discount rate.[43] The value for the total factor productivity parameter A is inconsequential, except for the scale of the variables in levels—with $A = 2$, most macroeconomic aggregates take two-digit values in levels, making them just easier to report.

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

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

corporate exposures in the standardized approach of Basel II.[44]

The values of p_0, p_1, and ε are set so as to have sufficient potential room for risk shifting and for significant aggregate losses due to it. The current choices are compatible with the conditions $(1-\varepsilon)p_1 + \varepsilon > p_0 > p_1$ established in assumptions A1 and A2. They imply unconditional *expected* default rates in the range from 3% (if all firms were non-systemic) to 4.7% (if all firms were systemic). The probability of a systemic shock (situation in which all systemic firms fail simultaneously) is set at 3%, so that this shock occurs on average once in every 33 years. These parameter choices help us extract our quantitative results in a context in which the efficiency losses due to risk-shifting are significant, while bankers' inclination towards risk-shifting is potentially large.[45]

The choices for the values of the bank capital dynamics parameters, ψ and ϕ, are quite tentative. The rate ψ at which bankers convert back into workers and stop providing their accumulated wealth as bank capital implies that they have an average active life of 5 years. The value of ϕ implies that bankers' represent 5% of the efficiency units of labor in the economy, so that their yearly labor income provides additions to the aggregate wealth available to active bankers equivalent to 5% of total labor income per period.[46]

5.2 Graphical presentation of the results

The numerical method used for solving the model relies on value function iteration in order to obtain $v(e)$. As described in Appendix C, we use the relevant equilibrium conditions in their full, potentially non-linear form.[47] All the results correspond to the equilibrium conditions described there, which consistently accommodate the possibilities of bankers wanting to

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

[45] The sensitivity of our results to changes in some of the key parameters is analyzed in Section 6.

[46] In the pseudo-steady state, the increase in active bankers' wealth due to ϕ exactly compensates the wealth taken out by the fraction ψ of non-continuing bankers, making the effective supply of bank capital constant until a systemic shock occurs.

[47] 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.

consume or save part or all of their wealth as bank deposits.[48] No problem of multiplicity of equilibria was detected.

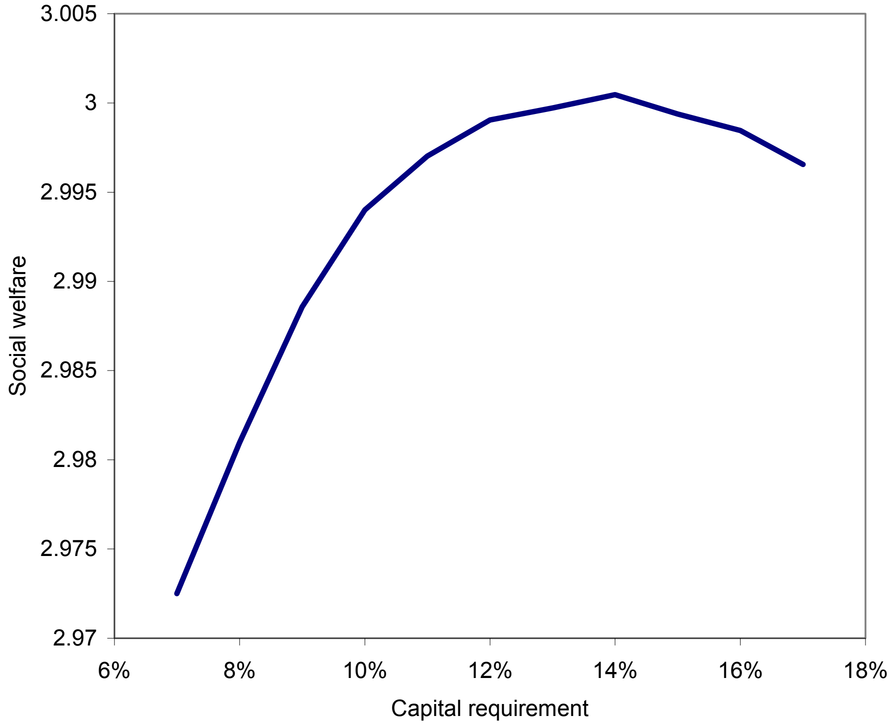


Figure 1 Social welfare W as a function of the capital requirement γ

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 $E(W_t)$.

For illustration purposes, we will henceforth compare the results obtained with a capital requirement of 7% (γ=0.07) with those obtained with a requirement of 14% (γ=0.14), which is the one that maximizes social welfare under our baseline parameterization. For these two requirements, Figure 2 depicts, for the two compared values of γ, the functions that describe the marginal value of one unit of bank capital $v(e)$ (top panel) and the fraction of bank

[48] Situations in which bankers do not want to fully reinvest their wealth as bank equity are typically confined to values of e out of the support of the ergodic distribution.

capital devoted to make systemic loans $x(e)$ (bottom panel). Both functions are depicted for a range of values of e that includes the union of the ranges relevant under each of the compared capital requirements. As later shown in Figure 3, total bank capital e fluctuates over a wider but lower range under a low γ than under the optimal γ.

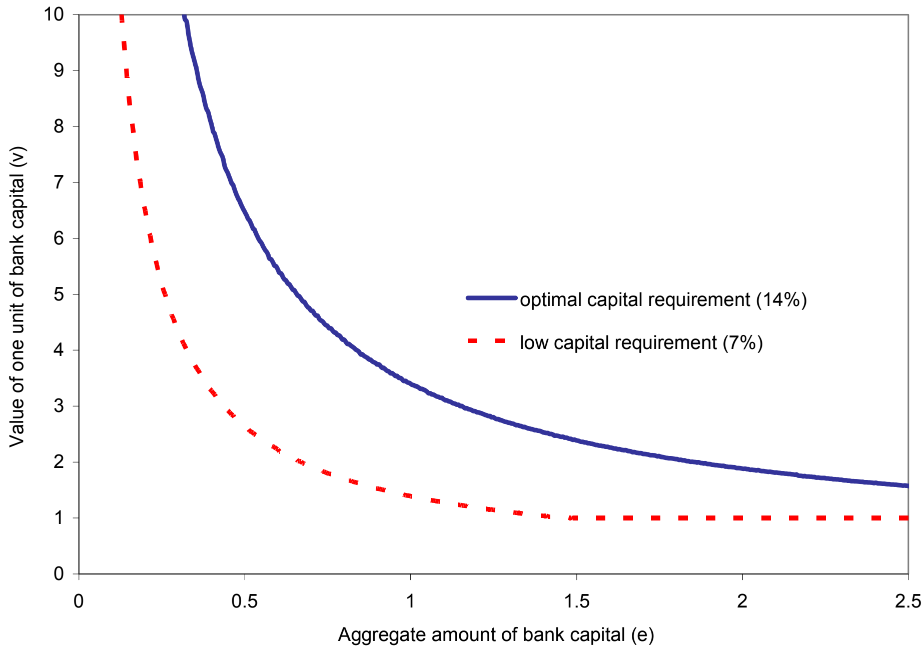
Figure 2 evidences that the greater scarcity of bank capital induced by a higher capital requirement implies a higher marginal value of capital $v(e)$ at every level of capital e. More importantly, systemic risk-taking, which is non-decreasing in e (in fact, strictly increasing in e up to the point in which bankers stop to fully reinvest their wealth as bank equity), is lower at every e with $\gamma = 0.14$ than with $\gamma = 0.07$.

Raising γ shifts $x(e)$ down in all the range of capital requirements that we have tried. The intuition for this result is that a higher γ implies higher scarcity and, thus, a higher equilibrium value of bank capital, $v(e)$, increasing the incentives for bankers to guarantee that their capital survives a systemic shock. This effect resembles the traditional charter value effect stressed in the microeconomic banking literature but, differently from it, has a general equilibrium origin (the impact of the systemic shock on the scarcity of bank capital), holds in a context with perfectly competitive banks, and is unambiguously reinforced when γ increases.[49]

The interaction between bankers' systemic risk taking, as reflected in $x(e)$, 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 lines in the top panels represent 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 $e_{t+1}^{1-\varepsilon}$. This schedule is strictly increasing except when e_t is large enough for continuing bankers to consume a part as a voluntary dividend (an option considered in Appendix C). At that point, the schedule becomes flat.

[49] The possible negative impact of capital requirements on risk taking identified by, for example, Hellmman at al. (2000) is due to the negative impact of the requirements on banks' profitability. Such a negative effect is established in a partial equilibrium setup with a perfectly elastic supply of bank capital and a finite number of imperfectly competitive banks.



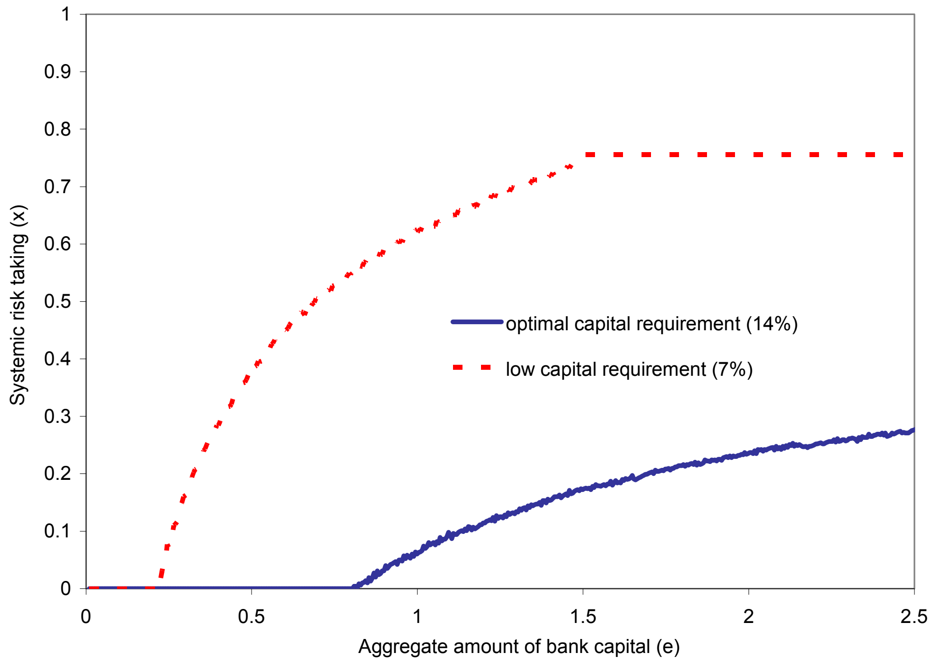


Figure 2 $v(e)$ and $x(e)$ under low and optimal capital requirements (CR)

The dashed downward sloping curve in the top panels represent the mapping from e_t to the capital available to bankers in the next date if the systemic shock occurs, e_{t+1}^{ε}. 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 (see Figure 2). Indeed, for sufficiently low values of e, we have $x(e) = 0$, in which case the two curves merge ($e_{t+1}^{1-\varepsilon} = e_{t+1}^{\varepsilon}$).

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 we have $x'(e) > 0$). 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. 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.[50] With $\gamma = 0.07$ ($\gamma = 0.14$) our economy fully recovers in a minimum of 5 (8) years.[51]

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 little frequent, our economy spends most of the time (about 80% or more) in the pseudo-steady state. However, when the systemic shock realizes the consequences are sizeable. Other points with positive frequencies correspond to the various recovery paths that may occur.

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

[51] 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.

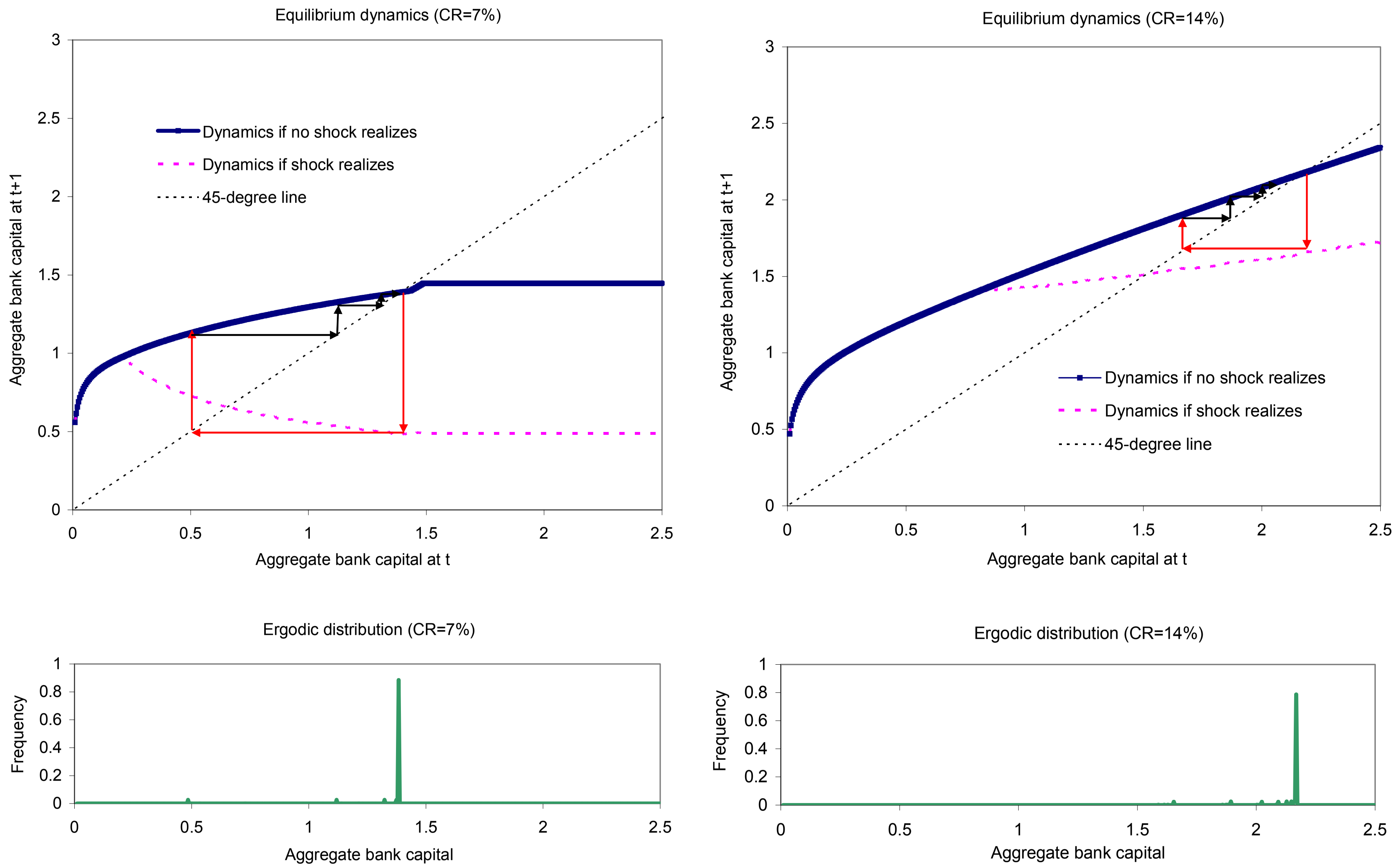


Figure 3 Equilibrium dynamics with low (CR=7%) and optimal (CR=14%) capital requirements

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 amount of systemic risk taking $x(e)$, particularly at the PSS, tells about the importance of the unresolved risk-shifting problem. Risk shifting produces "static" losses due to the inefficiency of operating the production technology in a way that implies a larger unconditional failure rate (and the subsequent waste of resources). Risk shifting also produces "dynamic" losses due the loss of bank equity (and the subsequent loss in banks' lending capacity) that follows the realization of a systemic shock and causes the amplification and propagation of its effects over time.

Fixing a higher capital requirement implies lower systemic risk-taking but does not guarantee a larger social welfare. Capital requirements have a negative impact on banks' lending capacity. To see this in our results, note that $\gamma = 0.14$ is twice as high as $\gamma = 0.07$ but, as shown in Figure 3, the PSS level of bank equity with $\gamma = 0.14$ is not twice as large as with $\gamma = 0.07$. So the larger capital requirement is associated with lower aggregate bank credit (and lower economic activity), at least in the PSS. This is a major counterbalancing effect to the beneficial welfare effect of capital requirements along the risk-taking dimension.

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 the list is social welfare, reported as $(1 - \beta)E(W_t)$, like in Figure 1. We find that the difference between the welfare associated $\gamma = 0.14$ 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 is the lower average fraction of bank capital invested in systemic banks (or devoted to support systemic loans): 24% rather than 71%.

Table 2 also shows that with a capital requirement of 14% (rather than 7%) bank loans are more expensive (with an unconditional mean loan rate of 5.4% rather than 4.1%). This implies obtaining significantly lower values in macroeconomic aggregates such as GDP and bank credit, whose unconditional means are, respectively, 6.5% and 20.7% lower with $\gamma = 0.14$ than with $\gamma = 0.07$.

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	0.94
GDP*	4.404	4.116	-6.52
Bank credit (l)	19.243	15.254	-20.73
Bank equity (e)	1.347	2.136	58.54
Loan rate (r_L) (in %)**	4.1	5.6	+1.5
Deposit insurance costs*	0.158	0.037	-76.54
Value of one unit of bank capital (v)	1.107	1.786	61.31
Fraction of capital invested in systemic banks (x)	0.706	0.244	-65.38

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

The analysis evidences that bank credit or 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.

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 in the period that follows a systemic shock. We compare their value one year after the shock if the economy gets hit at the PSS with the value in the PSS if no shock occurs. The shock wipes out a fraction of bank capital and leads to a contraction in the supply of credit. The increase in loan rates following the systemic shock is of 11.6 percentage points with capital requirements of 7% and of only 2.5 percentage points with capital requirements of 14%. Aggregate net consumption, GDP, and bank credit fall by 17%, 32%, and 65% with capital requirements of 7% and by only 5%, 10%, and 24% with capital requirements of 14%.

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*	-17.28	-4.55
GDP*	-31.74	-9.54
Bank credit (l)	-65.34	-23.96
Bank equity (e)	-65.34	-23.96
Loan rate (r_L) (difference in percentage points)	+11.6	+2.5
Value of one unit of bank capital (v)	160.29	25.79
Fraction of capital invested in systemic banks (x)	-50.07	-20.24

*See Appendix B for exact definitions of these variables.

Finally, Table 4 describes the values of some macroeconomic and financial ratios across the economies with $\gamma = 0.07$ and $\gamma = 0.14$, respectively. Some ratios show that our calibration is consistent with standard macroeconomic calibration targets. Other ratios point to weak capital regulation (low γ) as a potential cause of *financial exuberance*. In fact, the ratio of credit to GDP 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 a macroprudential approach to bank regulation (e.g. Borio, 2003).

Table 4
Other macroeconomic and financial ratios
(Unconditional expected values)

	$\gamma = 7\%$	$\gamma = 14\%$
Labor income/GDP*	0.69	0.68
Physical capital/GDP	3.68	3.03
Bank credit/GDP	4.37	3.71
Deposit insurance costs/GDP	0.036	0.009
ROE at non-systemic banks	1.104	1.168
ROE at systemic banks if systemic shock does not realize	1.189	1.213
Banks' payout/Bank capital	0.213	0.224

*See Appendix B for an exact definition of GDP.

All in all 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, give wrong indications about the desirability of setting capital requirements at one level or another. As we have seen, the optimal capital requirement of 14% seems to have a "large cost" relative to the suboptimal requirement of 7% when evaluated in terms of these variables.

6 Extensions and discussion

This section is structured in four parts. First, we analyze the transition from a regime with a low capital requirement to one with a higher target capital requirement, and assess the value of gradualism in approaching the 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 the effects of 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 main quantitative results to changes in some of the parameters.

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 that corresponds to $\gamma^0 = 0.07$. 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 $t = 0$ (when the capital requirement is 7%) 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.[52]

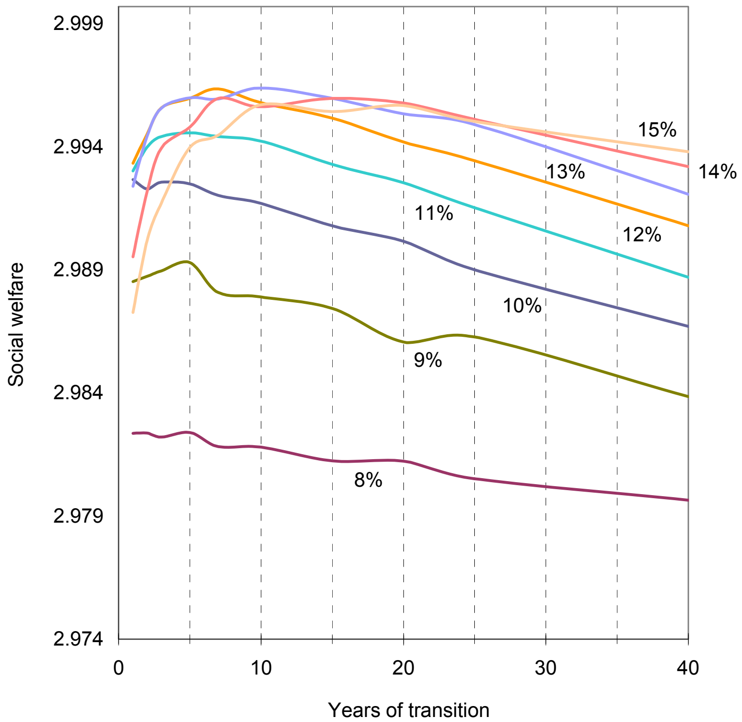


Figure 4 Social welfare for different target capital requirements as a function of the length of the transition (starting from the PSS with $\gamma = 0.07$)

Under our baseline calibration, the maximum welfare is obtained with a target capital requirement of 13% and 10 years of transition (although welfare is very similar with a target requirement of 12% and 7 years of transition). Interestingly, if capital requirements were to be increased only up to a (suboptimal) level of 10%, then the (conditional) optimal transition

[52]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 1000 equilibrium paths of 1000 years each starting from the PSS of the economy with $\gamma = 0.07$ and compute the welfare associated with each path.

length would just be $T = 1$, i.e. there would be no net gains from increasing the requirements in a gradual way. For higher target requirements, however, maximizing (conditional) welfare requires transitional periods whose lengths increase with the target.

Figure 4 also evidences the decreasing marginal social gains from increasing capital requirements in our economy—even when allowing for their gradual introduction. While there are significant gains from increasing capital requirements from 7% to up to 10% (even without gradualism), realizing significant gains from raising the target to, e.g., 11% (rather than 10%) requires gradualism (4 years). And the target of 12% (13%) only improves over the target of 11% (12%) if the transition is extended to 7 (10) years.

6.2 Cyclically-adjusted capital requirements

The model can also be extended 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 having $g' > 0$: it would contract (expand) the amount of credit that can be extended per unit of available bank capital when bank capital e_t is more abundant (scarcer) and equilibrium credit is likely to be excessive (insufficient).[53] By the contrary, making capital requirements tighter when bank capital is scarcer, $g' < 0$, would have undesirable procyclical effects.[54]

To check the effects of cyclically-adjusted capital requirements in a quantitatively easy manner, we consider the following flexible functional form for g:

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

where g_0 and g_1 are constant parameters, and $\overline{e}$ is the amount of bank capital that the economy would accumulate in its PSS under a time-invariant capital requirement $\gamma_t = g_0$.

[53] Gersbach and Rochet (2012) and Malherbe (2012) identify setups where contercyclical bank capital requirements increase efficiency in the intertemporal allocation of investment.

[54] Repullo and Suarez (2012) describe a setup where capital requirements (specially when they are risk-sensitive) have procyclical effects, but their welfare analysis identifies rather small net gains from the introduction of countercyclical adjustments.

Accordingly, the 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 cyclical adjustments would make more sense under our baseline parameterization of the model, we set $g_0 = 0.14$ (since a requirement of 14% maximizes social welfare among the class of time-invariant requirements) 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, up to a value of about $g_1 = -0.8$ (where the welfare gains relative to the time-invariant benchmark are of about 0.75%). In contrast, welfare falls quite sharply if g_1 is increased (with a welfare loss of about 2.5% relative to the benchmark for $g_1 = 0.2$). This implies that trying to produce countercyclical effects by relaxing capital requirements after a crisis is actually *counterproductive* in our model. If anything, there would be social welfare gains from making capital effectively scarcer for low values of e.

These results come from the endogeneity of systemic risk taking in the model. Bankers face a dynamic trade-off when choosing between systemic and non-systemic lending. One motivation for bankers to stay away from systemic risk taking is the expectation of earning scarcity rents on their surviving capital if the systemic shock realizes and other bankers lose their capital. Having $g_1 > 0$ would reduce the after-shock gains of the owners of surviving capital, pushing more bankers into inefficient systemic risk taking and increasing the severity of any systemic crisis. In other words, although $g_1 > 0$ reduces the credit crunch that follows any *given* loss of aggregate bank capital e_t, it also encourages ex ante risk taking by banks, implying that the endogenous size of bank capital losses is higher with $g_1 > 0$ than with $g_1 \leq 0$.

6.3 Bailout policies

What does our model say about bailout policies? Are they at all desirable? Do they have undesirable side effects? A detailed quantitative assessment of the many forms that these policies might take is beyond the scope of the current paper, but it is worth making a few remarks directly emanated from the logic of the model.

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. The tool to expand credit supply after the shock would in this case be transferring wealth to bankers (rather than reducing the capital requirement). 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.

To structure the discussion, 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. But in addition to the direct effects on credit supply there are two important incentive effects to consider.

First, the wealth transferred to the non-retiring 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 become less steep, producing a negative impact on incentives (i.e. increasing systemic risk taking as represented by $x(e)$). Second, there will also be direct incentive effects associated with bankers' prospects of receiving wealth transfers after the systemic shock—what many discussions about bank bailouts refer to 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 potentially most welfare enhancing manner of implementing a bailout would be by directing the wealth transfers exclusively to (non-retiring) bankers with wealth invested in banks that remain solvent after 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, offsets the negative welfare implications of the indirect incentive effects due to 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 shows, as a different curve, the effects 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% which is optimal in our baseline parameterization. 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).

7 Concluding remarks

The paper describes a dynamic general equilibrium model of endogenous systemic risk-taking. The model incorporates a meaningful definition of systemic risk and allows us to provide a formal assessment of the macroprudential role of bank capital requirements in a setup where some banks fail in equilibrium. Importantly, our normative analysis is performed using an internally consistent welfare measure, based on adding up the net consumption flows of the various agents in the model.

Our results suggest that capital requirements have qualitatively and quantitatively important effects on systemic risk-taking, standard macroeconomic variables, banking indicators, and welfare. Under our calibration of the model, socially optimal capital requirements are quite high, have a sizeable negative impact on GDP, and should be gradually introduced.

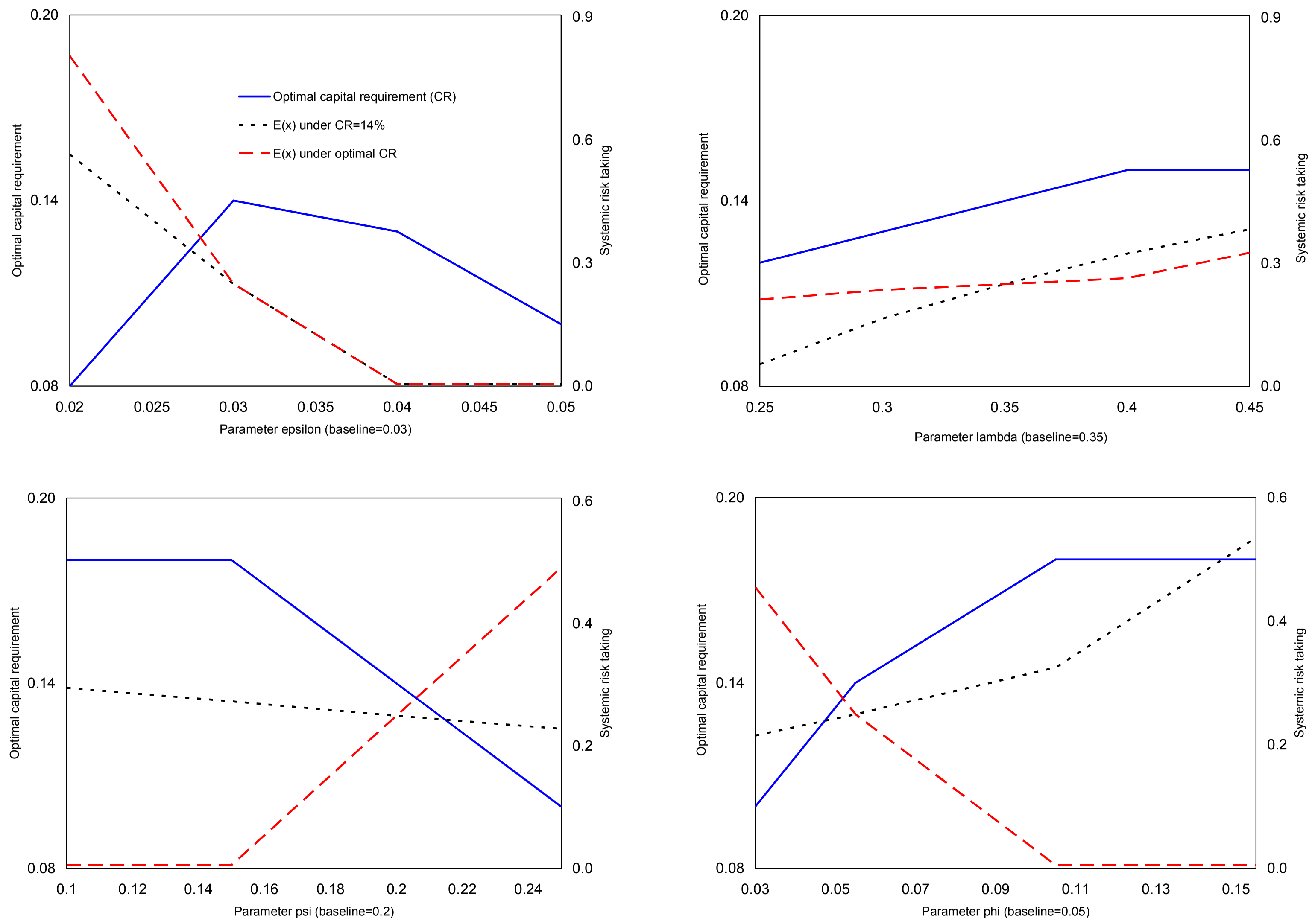


Figure 5 Dependence of the optimal γ (left axis) and systemic risk taking $E(x)$ (right axis) on some of the parameters

We also find that counter-cyclical adjustments to these requirements would be welfare decreasing and argue that, in the context of the model, bank bailouts may have negative effects on systemic risk taking and are most likely to be welfare enhancing if they concentrate the transfers of wealth on the owners of the banks that remain solvent after 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), $(1-p_0)B_t + p_0(1-\lambda)k_t - (1+r)d_t$. 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 the first constraint in problem (6), leaving bankers' participation constraint as follows:

$$E(v_{t+1})[(1-p_0)B_t + p_0(1-\lambda)k_t - (1+r)d_t] \geq Q_t e_t. \tag{18}$$

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, for what we just said, the return on equity at the non-systemic bank is not random. Hence (18) can be simplified to:

$$(1-p_0)B_t + p_0(1-\lambda)k_t - (1+r)d_t \geq R_{0t+1}e_t. \tag{19}$$

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 $(k_t, n_t, l_t, B_t, d_t, e_t)$.[55]

In the optimization problem stated in (6), both (19) and the constraint associated with the minimum capital requirement will be binding. 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:

$$B_t = \frac{1}{1-p_0}\{[(1-\gamma_t)(1+r) + \gamma_t R_{0t+1}](k_t + w_t n_t) - p_0(1-\lambda)k_t\}. \tag{20}$$

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

[55] 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.

date t. The term $p_0(1-\lambda)k_t$ 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

$$\max_{(k_t,n_t)} \ (1–p_0)[AF(k_t,n_t)+(1–\delta)k_t]+p_0(1–\lambda)k_t-[(1–\gamma_t)(1+r)+\gamma_t R_{0t}](k_t+w_t n_t), \quad (21)$$

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 FOCs 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.[56] 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 $1 + r_{Lt} = B_t/l_t$, where B_t is found by evaluating (20) at $n_t = 1$ and $l_t = k_t + w_t$.■

B Social welfare

The patient agents who provide (insured) deposit funding to the banks at rate r break even in terms of their own net present value in all periods. Thus their net consumption flows as depositors make a zero net addition to social welfare W_t and we can safely leave them out in the derivations that follow. All other agents have a discount factor $\beta < 1/(1 + r)$, 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).[57] Focusing on the case in which bankers always reinvest all their accumulated wealth as bank capital, social welfare

[56] 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 lending contract of the non-systemic bank.

[57] By assigning the cost of deposit insurance to the impatient agents, we prevent the possibility of increasing our utilitarian measure of social welfare by using deposit insurance as an indirect means of redistributing wealth from the patient agents to the impatient agents.

at any period t, W_t, can be expressed as

$$W_t = E_t \left(\sum_{s=0}^{\infty} \beta^s \omega_{t+s} \right), \tag{22}$$

where

$$\omega_t = -e_t + [1 - \phi(1+\psi)]w_t + \beta\{y_{t+1} - (1+r)[d_t - \phi(1+\psi)w_t]\}, \tag{23}$$

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

$$gdp_{t+1} = [(1-x_t)(1-p_0) + x_t(1-\varepsilon_{t+1})(1-p_1)]AF(k_t, 1), \tag{24}$$

$$\Delta_{t+1} = \delta + \{(1-x_t)p_0 + x_t[(1-\varepsilon_{t+1})p_1 + \varepsilon_{t+1}]\}(\lambda - \delta), \tag{25}$$

and $\varepsilon_{t+1} \in \{0, 1\}$ indicates whether the systemic shock realizes ($\varepsilon_{t+1} = 1$) or not ($\varepsilon_{t+1} = 0$) at the end of period t. In this decomposition, ω_t is the present value of the net consumption flows that the impatient agents derive from the 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 $t+1$, $\phi\psi w_t$, 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 (23)), the impatient agents in the economy (including the taxpayers who pay, if needed, for the net costs associated with deposit insurance) appropriate (if positive) or contribute (if negative) the difference between gross output y_{t+1} and the gross repayments $(1+r)[d_t - \phi(1+\psi)w_t]$ to the patient agents who hold bank deposits. Gross output is the sum of GDP as conventionally defined, gdp_{t+1}, and depreciated physical capital, $(1 - \Delta_{t+1})k_t$. The expressions for these two components, (24) and (25), 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 ε_{t+1}.

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

1. Impatient agents other than active bankers and next-period bankers, who receive wages at t and consume: $+[1 - \phi(1+\psi)]w_t$.

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

3. Entrepreneurs, who in such task break even state-by-state: $+0$.

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

5. Bankers 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, receive at $t+1$ the proceeds from having invested the wages earned at date t in bank deposits: $+\beta(1+r)\phi(1+\psi)w_t$.

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

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

C 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 $\gamma_i = \Gamma(e_i)$, where $\Gamma(\cdot)$ 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 conditions (8)-(10):

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

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

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

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

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

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

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

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

5. Consider the candidate $\{v_i\} = \{v(e_i)\}$. As an initial guess for $\{v(e_i)\}$, take some positive, non-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 $R_{1i}^{1-\varepsilon}$ for each i :

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

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

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

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

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

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

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

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

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

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

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

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

References

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

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

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

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, 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 (2011), "A macroeconomic model with a financial sector," mimeo, Princeton University.

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

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

Darracq Pariès, Matthieu, Christoffer Kok Sørensen, and Diego Rodriguez-Palenzuela (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.

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," CEPR DP 8764, January.

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 (2010), "Financial crises, bank risk exposure and government financial policy," *Journal of Monetary Economics*, forthcoming.

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

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

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

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 (2012), "Dynamic Macroprudential Regulation," 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 (2012), "The procyclical effects of bank capital regulation," CEMFI, mimeo.

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.