Optimal Capital Requirements over the Business and Financial Cycles*

Frederic Malherbe (University College London)

May 14, 2019

Abstract

I study economies where banks do not fully internalize the social costs of their lending decisions, which leads to real over-investment. The bank capital requirement that restores investment efficiency varies over time. During booms, more investment is desirable, so the banking sector must be allowed to expand. This suggests a loosening of the requirement. However, there also is more bank capital. Since the banking sector exhibits decreasing returns to scale, this suggests a tightening instead. I find that the latter effect, which I dub the *bank capital channel*, dominates: The optimal capital requirement is tighter during booms than in recessions.

1 Introduction

It is widely acknowledged that financial institutions have incentives to take on too much risk (Kareken and Wallace (1978), Acharya and Richardson (2009)) and that the banking crises that can result are costly for society.[1] Until the 2007-2009 global financial crisis, the main focus of

*I am grateful to Xavier Freixas, Anton Korinek, Marcela Lucchetta, David Martinez-Miera, Marcus Opp, Jaume Ventura, Lucy White, and Tanju Yorulmazer for their insightful discussions of the paper, to Saleem Bahaj, Patrick Bolton, Nuno Coimbra, Martin Eichenbaum, Emmanuel Farhi, Alessandro Graniero, Felipe Iachan, Enrique Mendoza, Matthias Paustian, Romain Ranciere, Helene Rey, Jean-Charles Rochet, Oren Sussman, Wolf Wagner, and seminar and conference participants at or in Amsterdam, Antwerp, the Bank of England, the Bank of Portugal, the Banque de France, Boston University, Cambridge, CREI, ECB, Exeter, Federal Reserve Board, FGV Rio, FGV Sao Paulo, Frankfurt, Georgetown, Gerzensee, HSE Moscow, IMF, Konstanz, Lancaster, KU Leuven, London Business School, London School of Economics, Manchester, the National Bank of Belgium, the New York Fed, Queen Mary, San Diego, the Sveriges Riksbank, Toulouse, Tilburg, Universitat Pompeu Fabra, Warsaw, Warwick, York, and Zurich for their comments and suggestions. This paper has been prepared by the author under the Lamfalussy Fellowship Program sponsored by the ECB. Any views expressed are only those of the author and do not necessarily represent the views of the ECB or the Eurosystem.

[1]For instance, the Savings and Loan crisis cost the US taxpayers at least \$132bn (in 1995 USD). There is no consensus on the 2007-2009 crisis' net fiscal costs, but Laeven and Valencia (2012) estimate the net outlays at 2.1% of GDP in the US, 6.6% in the UK, and up to a vertiginous 40% in Ireland. Besides fiscal costs, banking crises appear to severely affect the real economy as they are typically followed by long and painful recessions (Jordà et al. (2013)) and are associated with large permanent output losses (Cerra and Saxena (2008)).

bank prudential regulation was to contain risk taking at the individual bank level. Ever since, the focus has evolved towards containing system-wide risk taking. What are the costs and benefits of such policies? What are the relevant general equilibrium effects? Does an optimal policy vary over time?

To study these questions, I build a model of intertwined business and financial cycles. The model involves overlapping generations of risk-neutral savers and bankers. Bankers are protected by limited liability. They collect deposits and competitively lend to firms. Some firms succeed, some fail. Those that fail default on their bank loan. Banks are perfectly diversified, and the realized return to lending depends on the firm default rate. When the proceeds from lending are insufficient to repay depositors in full, the bank is insolvent and defaults.

Regulation is necessary because banks do not fully internalize the social costs associated with their lending decisions. This occurs for two main reasons: First, in the case of default, some of the costs are ultimately borne by the taxpayer. Second, credit expansion by a bank decreases the return to lending for other banks and therefore affects their probability of default and the associated deadweight losses. Because of these frictions, there is a market failure: banks lend too much, and real investment by firms is excessive.

The fist step of the analysis is to show that a regulator can restore investment efficiency thanks to a capital requirement. This optimal capital requirement is the ratio of aggregate bank capital (e) to the efficient level of aggregate lending (b^*). These two key macroeconomic variables evolve with the state of the economy, and their joint dynamics therefore pin down the cyclical properties of the optimal capital requirement.

Before turning to dynamics, I study the static determinants of b^*. Firms operate a constant return-to-scale technology. However, labor supply is inelastic, which implies diminishing returns to physical capital. This standard general equilibrium effect has important implications for the analysis.

The first is that, given aggregate bank capital, lending presents diminishing returns. So for a given e, there is an optimal b^*. Bank capital acts as a buffer to alleviate default costs. As a result, b^* is increasing in e.

Second, even though expected default costs are proportional to scale at the bank level, they are convex when we consider the banking sector as a whole. This is because the equilibrium marginal productivity of capital affects banks' expected default costs. As a result, the banking sector as a whole exhibits decreasing returns to scale. That is, doubling both b and e, the two inputs through which the banking sector generates economic surplus, less than doubles such surplus.[2]

Putting these insights together, we see that when aggregate bank capital goes up, lending should be allowed to expand, but *less than proportionally*. Since the capital requirement is binding in equilibrium, a proportional expansion would occur if the capital requirement was

[2] Absent default costs, this result trivially follows from diminishing returns to lending, as e does not play any role in alleviating default costs and therefore does affect economic surplus.

left unchanged. Instead, it should be tightened. Put simply: The optimal capital requirement is increasing in aggregate bank capital. This will play an important role in the results.

Productivity shocks are what trigger cyclical fluctuations in the model. They affect the optimal capital requirement through two main channels. First, a positive productivity shock raises expected productivity, which calls for additional investment. More aggregate lending is therefore needed. Keeping aggregate bank capital constant, increasing investment necessitates a *loosening* of the requirement. I dub this first mechanism, by which productivity shocks affect the optimal capital requirement, the *expected productivity channel.*

Second, a positive productivity shock raises aggregate bank capital. This is because improved productivity raises bank profitability and increases the inflow of wealth in the banking sector. Hence, even keeping expected productivity constant, the banking sector should be allowed to expand. But, as I have explained, this expansion should be less than proportional to the increase in e, and this calls for a *tightening* of the requirement. I dub this second mechanism, by which productivity shocks affect the optimal capital requirement, the aggregate *bank capital channel*.

I am able to isolate these two channels thanks to special cases that can be solved analytically. First, if productivity shocks are *iid*, only the bank capital channel operates. A positive shock must therefore be met by a tightening of the requirement. Second, if shocks are persistent, the expected productivity channel also operates. And the more persistent the shock, the stronger this channel. But persistent positive shocks also raise bank capital persistently. In fact, the effect of persistence is typically stronger for the bank capital channel than for the expected productivity channel.

As a result, positive productivity shocks generates a persistent tightening of the optimal capital requirement (and vice versa for negative shocks). So the optimal capital requirement is tighter in good times than in bad.

I prove this result analytically for the case in which default costs are negligible. Then I use numerical methods to confirm that it generally holds, and I calibrate the model to assess its quantitative relevance. First, I consider a productivity shock that implies a decrease in the firm default rate from 4% to 3%. The impulse response of the capital requirement is positive. Starting at a calibrated value of 8%, the capital requirement jumps to 8.1%, peaks at 8.4% after a few periods, and then slowly reverts to 8%. Second, I simulate periods of good times (i.e., when the loan default rate is at least one standard deviation below its mean) and bad times (when the default rate is at least one standard deviation above its mean). I find that the optimal capital requirement is substantially higher in the former situation (8.8% on average) than in the latter (5.8%).[3]

I also compare the optimal capital requirement to a suboptimal policy that takes into account the expected productivity channel but ignores the bank capital channel. Under this suboptimal

[3]For the sake of concreteness, I express these numbers in terms of a requirement that applies to risk-weighted assets, where risk weights are calibrated to 50%.

policy, the cyclical properties of the requirement are reversed: It is tighter in bad times than in good. As a result, banks take on too much risk in good times (they lend too much and retain less profit), and credit contraction is unnecessarily severe in bad times. According to my calibration, the associated welfare losses amount to 0.04% of steady state consumption, which aligns with numbers found in the literature on stabilization policy (Lucas, 2003).

The features of the suboptimal policy resonate with criticisms of the second regime of the international standards for banking regulation (BCBS (2004)), commonly referred to as *Basel II*. The typical criticism focuses on the pro-cyclical effect of risk weights. Risk weights are linked to loan default probabilities, which are lower in booms than in recessions. As a result, capital requirements are effectively looser in good times than in bad times. The consensus is that the overall effect of this magnifies the credit cycle. To mitigate such mechanisms, Basel III, the the current regulatory regime (BCBS (2010)), introduced the so-called countercyclical capital buffers (typically abbreviated as CCyBs).

Taking the cyclical effect of risk weights as given, I use my calibration results to compute back-of-the-envelope estimates for the time-varying adjustments to the optimal capital requirement. I find that a 1% decrease in the default rate for firms should be met, at the 4-year horizon, by an increase of the requirement from 8% to 9.12%. This can be interpreted as a 112bps CCyB. The 112bps correspond to 80bps needed to neutralize the decrease in risk-weights, plus 78bps to account for the bank capital channel, minus 56bps to account for the improved credit quality (which reflects the expected productivity channel). This suggests that accounting for the bank capital channel may be quantitatively as important as accounting for the procyclical effects of risk weights. It also suggests that the welfare gains from financial stabilization policy could be twice as large as those I find in the exercise described above.

Highlighting the bank capital channel, demonstrating how it arises from a simple general equilibrium effect, and showing that it is quantitatively relevant is the main contribution of this paper.

A key element of my analysis is that the regulator can restore efficiency thanks to capital requirements. In effect, capital requirements put restrictions on equilibrium quantities (i.e., aggregate lending). This pins down the equilibrium because the frictions of the model (deposit insurance and default costs) induce externalities that always generate excessive lending. However, other frictions in the banking sector –for instance market power– could lead to under-lending instead. In that case, capital requirements are unlikely to work as well, and other regulatory tools are typically needed.[4] In the final section of the paper, I discuss alternative modeling assumptions and study a variation of the model that can deliver under-investment. I show that, when this happens, efficiency can be restored with a Pigouvian approach, but in general, not with a capital requirement.

[4]In Gersbach and Rochet (2017), where under-lending is part of the problem, capital requirements that depend on the state of the economy improve welfare but do not fully restore efficiency. When under-lending arises due to overhang problems, whether capital requirements increase or reduce lending depends on the joint payoff distribution of existing and new loans (Bahaj and Malherbe (2018)).

Frictions that induce over-lending and over-investment seem particularly relevant to many countries when we consider the decade that led to the global financial crisis. Spain is a case in point. For instance, Santos (2014) reports that the credit boom generated an excess supply of hundreds of thousands of dwelling units. Diaz and Franjo (2015) also point out inefficiently high investment in structures and highlight the importance of government explicit and implicit subsidy in this process (see also See also Garcia-Santana et al. (2016) and Fernández-Villaverde et al. (2013)). During the same period in the US, banks were able to borrow at substantially subsidized rates thanks to government guarantees (Acharya et al. (2016); Atkeson et al. (2018)), and the economy also experienced a boom in real-estate investment (Rognlie et al. (2018)). Since the problem seems often related to real estate, one could interpret the optimal capital requirement of my model as a sectoral capital requirement, applied to real-estate lending. This is with the caveat, however, that my model cannot address how banks allocate their capital across sectors.[5]

This paper belongs to the literature on the costs and benefits of bank capital regulation. As explained above, it is is well understood that a regulatory regime like Basel II is likely to magnify the business cycle (Kashyap and Stein (2004), Repullo et al. (2010)). Adjusting capital requirements to the aggregate state of the economy seems a sensible response. However, the ways in which such adjustments should be designed remains an open questions.

Kashyap and Stein (2004) argue in favor of adjustments based on the scarcity of aggregate bank capital (relative to lending opportunity). They point out that it is not obvious *a priori* that aggregate bank capital scarcity is greater during a recession, but they interpret the empirical literature on bank capital crunches as generally supporting such a notion. My theoretical model delivers such a result.[6] Repullo and Suarez (2013), however, find the opposite. They study a model of optimal bank capital requirements and compare them to Basel I, II, and III. In their setup, even though bank capital is scarcer in bad times, which suggests adjustments that go in the direction of Basel III's CCyBs, capital requirements should still optimally be tighter in bad times than in good. However, they model a production function that is linear in physical capital. As a result, they do not capture the general equilibrium effect that drives the bank capital channel.

Martinez-Miera and Suarez (2014) propose a model where correlated risk-shifting by some banks gives other banks an incentive to play it safe. The reason is that banks that survive a crisis earn large scarcity rents in the aftermath, an application of the "last-bank-standing effect" (Perotti and Suarez (2002)). They mainly focus on the optimal *level* of a constant capital requirement.[7] However, they also consider a loosening of the capital requirements after

[5] A reason why over-investment is likely to materialize in real estate is that mortgages are easily securitized, and securitization can help regulation arbitrage, and gives opportunities to banks to magnify their systemic risk exposure (Malherbe and McMahon (2019)).

[6] In a static model, Repullo (2013) finds that capital requirements should be loosened after an exogenous negative shock to bank capital. This is in line with Kashyap and Stein (2004)s premise and with my results, but the mechanism is different.

[7] Other recent papers focusing on the optimal level include Morrison and White (2005), Van den Heuvel (2008),

a banking crisis. They find mixed effects for such a policy. It alleviates the credit crunch, but this mitigates rents ex post, which induces more systemic risk-taking ex ante. In contrast, Dewatripont and Tirole (2012) find that incentives to gamble for resurrection are stronger after a negative macroeconomic shock. Finally, in a recent paper, Collard et al. (2017) study how monetary and prudential policies can be jointly optimal.

The paper is organized as follows: I present the model in Section 2. I establish the market failure and derive the optimal regulatory response in Section 3. I focus on the general equilibrium effect in Section 4. I study the cyclical properties of the optimal capital requirement in Section 5. I provide a quantitative assessment in Section 6. In Section 7, I discuss alternative modeling choices. Then I conclude.

2 The baseline model

There are an infinite number of periods indexed by $t = 0, 1, 2...$, in which a single consumption good is produced and used as the unit of account, and where generations of agents born at different dates overlap.

Agents All agents are risk neutral, live two periods, and derive utility from their end-of-life consumption. There is a measure 1 of agents born at the beginning of each period. They are endowed with one unit of labor, which they supply inelastically during the first period of their lives.

After having worked and received their wage, a measure $\eta \ll 1$ of these agents become endowed with banking ability, which enables them to set up a bank under limited-liability protection. I refer to these agents as bankers. The remaining mass $1 - \eta$ of agents become savers.

Savers choose how to allocate their after-tax wages between deposits and a storage technology that provides a zero rate of return. I focus on cases where the storage technology is used in equilibrium, so that savers are indifferent between the two options.[8] Bankers have these options too, but they can also invest their wealth in their bank's equity.

Banks There is a continuum of banks, founded by bankers, that issue deposits to savers and lend competitively to firms. The rate of return to lending is denoted r. It is a random variable, whose distribution is determined in equilibrium, and taken as given by bankers.

When a bank's proceeds from lending are not sufficient to repay its deposits, the bank defaults. Bank defaults generate deadweight losses, which I model through a decrease in the recovery value of the bank assets. That is, in the event of default, shareholders receive nothing

Admati et al. (2010), and Begenau (2015).

[8] The economy can be considered as a small open economy with excess savings. Alternatively, I could simply allow banks to raise deposits from the rest of the world.

and the value recovered by depositors is lower than the value of the bank's assets. In practice, deadweight losses can arise for a number of reasons (e.g., costly state verification, distortions from taxation generated by an associated bailout, or other forms of spillovers to the real economy). Formally, I capture these losses with a default cost function Ψ that may depend on endogenous variables such as the bank equity (e), its level of lending (b), and the realized return to lending (r). This function satisfies the following regularity conditions: (i) $0 \leq \Psi < \infty$; (ii) $\forall \Psi > 0$: $0 \leq \Psi_b < \infty$, and $-\infty < \Psi_e \leq 0$.

In the case of bank default, the government compensates depositors for any loss they have suffered. Hence, deposits pay a zero interest rate. The government does not charge an insurance premium ex ante but, when needed, levies lump-sum taxes on savers in order to break even for that period.

Firms In each period, there is a continuum of penniless firms, indexed by i, that operate a constant-return-to-scale production function. Firms competitively hire labor and borrow from banks to invest in physical capital. For simplicity, firms operate for only one period, and physical capital can be transformed one for one into the consumption good and vice versa. The production function takes the form $a_i k_i^{\alpha} n_i^{1-\alpha}$, where n_i denotes labor, k_i is physical capital ($0 < \alpha < 1$), and $a_i \in \{0,1\}$ is a random variable that captures firm-specific productivity.

The realization of a_i is observed by firm i at no cost, but it is costly to observe for the banks. As is standard, I assume that such verification costs materialize as a reduction in the liquidation value of the firm. As a result, the optimal contract is a debt contract, whereby the firm repays the principal plus a given interest rate r^l when $a_i = 1$, and enters bankruptcy when $a_i = 0$. To formalize verification costs, I follow Martinez-Miera and Suarez (2014) and assume that while capital depreciates at a rate δ when $a_i = 1$, it depreciates at a larger rate $\delta + \Delta$ when $a_i = 0$ and the firm enters bankruptcy. Finally, I assume that failure is independent across firms and denote $A \equiv E[a_i]$ the commonly known probability that a given firm succeeds.

Shocks To trigger fluctuations in this economy, I let A evolve according to an AR(1) process:

$$\ln A_t = (1-\rho)\ln \bar{A} + \rho \ln A_{t-1} + \epsilon_t, \tag{1}$$

where ρ is a positive parameter capturing the persistence of the shocks, $\bar{A}$ is a positive constant, and ϵ_t represents an *iid* shock with mean zero and standard deviation $\sigma > 0$. The realization of A_t is the event that defines the beginning of period t.

Period timeline

1. A_t is realized and publicly observed, and firms competitively hire workers (firm capital stock is predetermined by decisions as date $t-1$).

2. Production takes place and is allocated: Wages are paid, solvent firms repay their loan, and insolvent firms transfer their residual value to the banks.
3. If solvent, banks repay depositors. If insolvent, banks default, and the associated costs are incurred by the depositors. The regulator compensates them for their losses and taxes savers to break even.
4. Young agents learn whether they have banking ability. Young bankers make their investment portfolio decisions (storage and/or investment in their bank's equity).
5. Banks borrow from savers and lend to firms, which invest in physical capital.
6. The older generation consumes and leaves the economy.

3 Market failure and constrained efficiency

3.1 Competitive equilibrium

The problem of the banker Bankers' relevant decisions are how to allocate their wealth between storage and bank equity and how much the bank should lend, given its level of equity. This can be formalized as follows.

Consider a representative bank at date t, and denote by e_t its amount of equity. To lend an amount b_t, the bank needs to raise $b_t - e_t$ of deposits. Let v_{t+1} denote the ex-post net worth of the bank–i.e., its value after its return to lending r_{t+1} is realized. That is,

$$v_{t+1} \equiv b_t r_{t+1} + e_t - \Psi_{t+1}$$

where $\Psi_{t+1} \geq 0$ is the realized default cost.

Then, consider a representative banker born at date t that has earned a wage w_t. He solves:

$$\max_{e_t, b_t} E_t \left[c_{t+1} \right] \tag{2}$$

subject to the budget constraints and non-negativity conditions:

$$\begin{cases} e_t + s_t = w_t \\ c_{t+1} = v_{t+1}^+ + s_t \\ e_t,\ b_t,\ s_t,\ c_{t+1} \geq 0 \quad , \end{cases} \tag{3}$$

where c_{t+1} represent his consumption, s_t is the amount he stores from date t to date $t+1$, and v_{t+1}^+ is the realized (private) value of his bank's equity–i.e., the positive part of v_{t+1}:

$$v_{t+1}^+ \equiv \left[b_t r_{t+1} + e_t \right]^+ . \tag{4}$$

Note that Ψ_{t+1} does not appear in (4), because when $b_t r_{t+1} + e_t \geq 0$ it is nil.

Labor market Labor is hired at the beginning of the period–that is, after A_t is known, but before a_{it}'s are realized. The labor market is competitive. The expected wage is:[9]

$$w_t = (1-\alpha)A_t k_t^{\alpha}$$

The return to lending At the end of a given period $t-1$, firms borrow competitively from banks to form capital that they will use for production in period t. Hence, investment takes place before A_t is known.

In a competitive equilibrium, the expected unit repayment to the bank equates the firms' expected marginal return to capital (accounting for verification cost). Since firms are penniless and protected by limited liability, there cannot be states in which they make strictly positive profits (otherwise they would make profits in expectation). As a result, in an optimal contract, the date-t repayment to the bank by a firm i that had borrowed k_{it} and hired n_{it} workers, must correspond to its realized net share of capital. That is:

$$\begin{cases} \alpha k_{it}^{\alpha} n_{it}^{1-\alpha} + (1-\delta)k_{it} & ;\ a_{it} = 1 \\ (1-\delta-\Delta)\,k_{it} & ;\ a_{it} = 0. \end{cases}$$

Date $t-1$ capital market clearing requires $k_t \equiv \int_i k_{it} = b_{t-1}$, and date t labor market clearing requires $n_t \equiv \int_i n_{it} = 1$. Given constant return to scale, all firms have the same capital-to-labor ratio in equilibrium. A standard debt contract with an interest rate $r_t^l \equiv \alpha k_t^{\alpha-1} - \delta$ is therefore optimal. Accordingly, the realized rate of return to lending for the bank corresponds to the average net marginal return to capital. That is:

$$r_t \equiv r(A_t, k_t) = \alpha A_t k_t^{\alpha-1} - (\delta + (1-A_t)\Delta)\,.$$

Competitive equilibrium definition Given a sequence for the random variables $\{A_t\}_{t=0}^{\infty}$, and initial condition k_0, a competitive equilibrium is a sequence $\{w_t, r_{t+1}^l, e_t, b_t, \tau_t\}_{t=0}^{\infty}$, such that: vector $\{w_t, r_{t+1}^l\}$ clears the labor and capital markets at date t; vector $\{e_t, b_t\}$ solves the maximization problem of the representative banker born at date t; and τ_t is a lump-sum tax on savers such that the regulator breaks even at all t.

3.2 Market failure and regulatory response

Investment efficiency Economic surplus in this economy corresponds to output, net of depreciation and bank default costs. Formally, economic surplus at date $t+1$ is given by:

[9] At failed firms, the realized wage is nil. This is unrealistic, but since agents are risk neutral, it is the expected wage that matters. An alternative would be to assume that wages are paid in advance of production (this would require the firm to borrow additional funds).

$$S(A_{t+1},k_{t+1},e_t) \equiv A_{t+1}k_{t+1}^{\alpha} - (\delta + (1-A_{t+1})\Delta))\,k_{t+1} - \Psi\,(A_{t+1},k_{t+1},e_t)\,. \tag{5}$$

Investment efficiency requires to maximize expected economic surplus. Define:

$$k_{t+1}^{*} \equiv \arg\max_{k_{t+1}} E_t\left[S(A_{t+1},k_{t+1},e_t)\right] \tag{6}$$

Definition 1. Investment at date t is efficient if and only if $k_{t+1} = k_{t+1}^{*}$.

I can now compare the competitive equilibrium capital stock, which I denote k_{t+1}^{CE}, with this efficiency benchmark (which I discuss in Section 7).

Proposition 1. *At all dates, the competitive equilibrium capital stock is inefficiently high. That is: $k_{t+1}^{CE} > k_{t+1}^{*}, \forall t$.*

Proof. All the proofs are in Appendix A. □

Because of deposit insurance, banks do not fully internalize the losses that occur in bad states. This reflects an implicit subsidy. Banks compete for lending to firms and pass this subsidy onto them in equilibrium. As a result, firms over-invest. In this context, mispriced deposit insurance is a sufficient ingredient to cause over-investment, but it is not necessary. This is because credit expansion by a bank also increases the probability of default of other banks and the losses they incur in such an event. As I explain in the next section, this is an externality that is not internalized in a competitive equilibrium.

The regulator I study the problem of a regulator, whose mission is to restore investment efficiency. The regulatory tool is a time-varying capital requirement $x_t \in [0,1]$ that constrains banks' lending to a multiple of their equity.

$$e_t \geq x_t b_t. \tag{7}$$

Henceforth, I refer to e_t as *bank capital* (whereas I often simply refer to physical capital (k_t) as *capital*).

Constrained equilibrium A *constrained equilibrium* is defined as a straightforward extension of the competitive equilibrium. Given the same sequence of random variables and initial condition, it is defined as a sequence of capital requirements $\{x_t\}_{t=0}^{\infty}$ and a vector sequence $\{w_t, r_{t+1}^{l}, e_t, b_t, \tau_t\}_{t=0}^{\infty}$ satisfying the same conditions, the only difference being that $\{e_t, b_t\}$ must solve the problem of the representative banker born at t *subject to the capital requirement x_t*.

The optimal capital requirement I restrict my analysis to the interesting case in which bank capital is *scarce* (i.e., $\eta_t w_t < k_{t+1}^{*}$) at all dates t. If bank capital was plentiful, the optimal capital requirement would be $x_t = 1$, in which case banks would be fully funded with equity and could not default in equilibrium.

Proposition 2. *The regulator can ensure investment efficiency at all t with the following capital requirement:*

$$x_t^* \equiv \frac{\eta_t w_t}{k_{t+1}^*}. \tag{8}$$

Corollary 1. *(Equilibrium characterization) In a constrained equilibrium, the capital requirement is binding, $e_t = \eta_t w_t$, $s_t = 0$, $b_t = k_{t+1}^*$, and the equilibrium value of w_t and r_t^l are pinned down by their respective market-clearing conditions.*

To understand the dynamics of the model, it is useful to express the optimal capital requirement as:

$$x_t^* = \frac{e_t}{b_t(A_t, e_t)},$$

and to study the determinants of b_t in a given period. This is what I do next.

4 Optimal lending and general equilibrium effect

Firms operate a constant return-to-scale technology. However, aggregate labor supply is not perfectly elastic (indeed, it is fixed), which means that there are diminishing returns to physical capital in the aggregate.

This completely standard general equilibrium effect has important implications for the analysis. The first one is that, given aggregate bank capital, lending presents diminishing returns. So, for a given e, there is an optimal b^*. Bank capital acts as a buffer, and therefore alleviates default costs. As a result, b^* is increasing in e.

Second, even though expected default costs are proportional to scale at the bank level, they are convex when we consider the banking sector as a whole. This is because the equilibrium marginal productivity of capital affects the probability of banks' default and their realized default costs. As a result, there are decreasing returns to scale in the banking sector.

Below, I establish these results formally and derive an important consequence: The optimal capital requirement is increasing in aggregate bank capital. For simplicity, I focus on a single period and assume full continuous support for A and that all the relevant derivatives exist. Additionally, I omit time subscripts for a better readability.

The planner's problem Since the optimal capital requirement is binding, the problem of the planner can be written directly in terms of aggregate lending:

$$\max_{b \geq 0} E[A]b^{\alpha} - (\delta + (1 - E[A])\Delta))\, b - \int_0^{A_0(b)} \Psi(A,b) f(A) dA. \tag{9}$$

where $A_0(b) \equiv (A \mid v(b) = 0)$.

4.1 The determinants of optimal lending b^* (given e)

The planner's first order condition is:

$$\underbrace{\alpha E[A] b^{\alpha-1} - (\delta + (1-E[A])\Delta))}_{=E[r]} - \underbrace{\overline{\Psi}_b}_{\text{marg. exp. default costs}} = 0, \tag{10}$$

where

$$\overline{\Psi}_b \equiv \underbrace{\frac{\partial A_0}{\partial b}}_{\text{change in PD}} f(A_0)\Psi(A_0,b) + \int_0^{A_0(b)} \underbrace{\frac{\partial \Psi(A,b)}{\partial b}}_{\text{change in RDC}} f(A)dA.$$

Decreasing returns to lending The first two terms in the first order condition correspond to the equilibrium expected return to lending for banks. Since $k = b$ in equilibrium, the first term makes it obvious that diminishing returns to physical capital imply diminishing returns to lending.

But as we will see, diminishing returns to capital also affect the marginal expected default costs $\overline{\Psi}_b$ through both the probability of default (PD) and the realized default costs (RDC).

Default costs function At this point, I need to specify a functional form for Ψ. I consider a simple default cost function where deadweight losses are proportional to the bank's financial distress–that is, they are proportional to the bank's (negative) net worth. Formally:

$$\Psi(b,k) \equiv \gamma\left[-(br(k)+e)\right]^+, \tag{11}$$

where γ is a positive and finite parameter; operator $[\ .\ ]^+$ selects the positive part of the argument; and $br + e$ is the realized net worth (where r is the random variable). Hence, by construction, the costs only occur when net worth is negative–that is, when the bank defaults.

There are several possible interpretations for such a specification. For now, to fix ideas, it is useful to think of deadweight losses as emanating from the bank resolution process, which is arguably longer and more complicated when financial distress is high. Besides being intuitively appealing, default cost function (11) is analytically convenient (Townsend, 1979). Note, however, that the main results do not hinge on this specification (see Section 7 for alternative specifications and a discussion).

Realized default costs in general equilibrium Assume that $br + e < 0$ and the bank defaults. Then, at the bank level, we have $\frac{\partial \Psi(b,k)}{\partial b} = -\gamma r$. Realized default costs are linear in b. However, in the aggregate, we need to take r's dependence on k into account. This makes realized default costs convex. Formally, since default requires $r < 0$, we have:

$$\frac{d\Psi(b,b)}{db} = -\gamma\left(r + br'(b)\right) > -\gamma r.$$

Probability of default in general equilibrium Default threshold A_0 is defined by:

$$\underbrace{b\left(\alpha A k^{\alpha-1} - (\delta + (1-A)\Delta)\right)}_{=r} + e = 0,$$

which gives:

$$A_0(b,k) \equiv \frac{(\delta+\Delta)\,b - e}{\left(\alpha k^{\alpha-1} + \Delta\right) b}.$$

We can see clearly that A_0 is increasing in k. Hence, an increase in aggregate lending increases a given bank's probability of default. This adds to the convexity of default costs in the aggregate.

Remark 1. Banks do not take the general equilibrium effect into account when making their decisions. As a result, they do not internalize that their level of lending affects the probability of default of other banks, and their costs given default. This is an externality that they do not internalize in equilibrium. Since the general equilibrium effect magnifies expected default costs, this means that the competitive equilibrium would exhibit over-lending even in the absence of deposit insurance.[10]

Remark 2. Since the capital requirement indirectly targets efficient quantities, the size of the wedge (between the planner's and the representative bank's first order conditions) created by this externality does not directly matter for the optimal requirement. Put differently, a higher optimal capital requirement does not necessarily reflect an increase in the magnitude of this wedge. Furthermore, the magnitude of the wedge does not determine b^*–it determines how much the competitive equilibrium deviates from b^*.

Optimal lending and expected productivity A desirable property of my specification for Ψ is that, for a given b, expected default costs are decreasing in expected productivity. This guarantees that optimal aggregate lending increases with expected productivity. The following lemma formalizes this in terms of elasticity.

Lemma 1. *Denote ξ the point elasticity of b^* with respect to $E[A]$. For all $E[A]$, $\xi(E[A]) > 0$.*

4.2 Optimal lending and bank capital

First order condition (10) implicitly defines a function $b^*(e)$. This function is increasing, which is intuitive.[11] An important point for the analysis is that changes in b^* are *less than proportional* to those in e. This is important because it implies that the optimal capital requirement is

[10]Deposit insurance implies a very simple externality: in the event of default, the negative net worth of the bank plus the default costs are shifted onto the taxpayer. The implicit subsidy that this entails also contributes to over-lending.

[11]This property also arises for fixed default cost functions, for costs proportional to the size of the balance sheet, for costs proportional to the ex-post value of the assets, and at least for any function that satisfies $\frac{\partial \Psi}{\partial e \partial b} \leq 0$.

increasing in aggregate bank capital. To see why, just write the optimal capital requirement as:

$$x^*(e) = \frac{e}{b^*(e)}.$$

Now, to understand why b^* increases less than proportionally to e, and how this is linked to the general equilibrium effect, first consider an atomistic bank. Its expected net worth (from a social point of view) is

$$v(b,e) \equiv bE[r] + e - \gamma \int_0^{A_0(b,e)} [-(br+e)]^+ f(A)dA,$$

where

$$A_0(b,e) \equiv \frac{(\delta+\Delta)\, b - e}{\left(\alpha k^{\alpha-1} + \Delta\right) b}.$$

Scale up this bank by a factor λ–that is, consider $v(\lambda b, \lambda e)$. Note that $A_0(\lambda b, \lambda e) = A_0(b,e)$–that is, the default threshold depends on leverage, not on scale. Given that realized default costs are linear in scale, this directly implies that $v(\lambda b, \lambda e) = \lambda v(b,e)$. The social value of an atomistic bank is proportional to its scale.

However, when we consider the banking sector as a whole, we get decreasing returns to scale. First, $E[r]$ decreases in k, but it is not affected by e. Second, even though an increase in e mitigates default costs (in a way that makes them linear in scale λ at the bank level), this is not enough to maintain constant returns to scale in general equilibrium. As above, this is due to the impact of diminishing returns to capital on the default threshold and on the realized default costs. Hence, the banking sector as a whole exhibits *decreasing returns to scale*. Together with concavity in b, this explains why $b^*(e)$ is increasing less than proportionally to e.[12] The following proposition formalizes this argument in terms of elasticity, and its corollary summarizes its main implication.

Proposition 3. *Denote κ the point elasticity of b with respect to e. For all $e \geq 0$, $0 \leq \kappa(e) < 1$.*

Corollary 2. *The optimal capital requirement is increasing in aggregate bank capital.*

[12]Technically, what ensures that the elasticity is smaller than one is decreasing returns to scale in the marginal social benefit of lending. See the proof for details.

5 Cyclical properties of the optimal capital requirement

The following system, together with (1), the law of motion for A_t, fully captures the dynamics of the model:[13]

$$\begin{cases} e_t &= \eta(1-\alpha)A_t\left(b^*_{t-1}\right)^\alpha \\ b^*_t &= b^*_t\left(E[A_t],e_t\right) \\ x^*_t &= e_t/b^*_t \end{cases} \tag{12}$$

In Section 6, I analyze the model globally. In this section, I derive analytical results on the basis of a linearized version.

5.1 Linearized model

Risky steady state Without risk in the economy, banks would never fail in equilibrium. Hence, there would be neither default costs nor transfers from the taxpayer. Without these distortions, the competitive equilibrium would be efficient, and there would be no need for a capital requirement. A deterministic steady state is therefore not a meaningful starting point from which to study fluctuations. Instead, I use the simple notion of risky steady state: the state of the economy that repeats itself if agents expect future risk and if the realization of shocks is zero (Coeurdacier et al., 2011). Henceforth, I use the subscript ss to denote variables at the risky steady state, which I simply refer to as the steady state. Formally, it is defined as the following fixed point:

$$\begin{cases} e_{ss} = & \overline{A}\eta(1-\alpha)\left(b^*_{ss}\right)^\alpha \\ b^*_{ss} = & b^*_{ss}\left(\overline{A},e_{ss}\right) \\ x^*_{ss} = & \frac{e_{ss}}{b^*_{ss}} \end{cases}$$

Intuitively, this corresponds to the regulator choosing the efficient level of aggregate lending given the risks facing the economy, but those risks never being realized. Solving numerically for the steady state is straightforward.

Elasticities Defining a_t as the log-deviation of A_t from $\overline{A}$, we have:

$$\begin{cases} \boldsymbol{e}_t &= a_t + \alpha\boldsymbol{b}_{t-1} \\ \boldsymbol{b}_t &= \xi E_t\left[a_{t+1}\right] + \kappa\boldsymbol{e}_t \\ \boldsymbol{x}_t &= \boldsymbol{e}_t - \boldsymbol{b}_t \end{cases} \tag{13}$$

[13]The first equation of the system is obtained by substituting the labor and capital market clearing conditions in the equilibrium expression for e_t (see Corollary 1).

where bold variables $\boldsymbol{e}_t$, $\boldsymbol{b}_t$, and $\boldsymbol{x}_t$ are the log deviation from e_{ss}, b^*_{ss}, and x^*_{ss}; and ξ and κ are the point elasticities of b^*_{ss} with respect to $E_t[A_{t+1}]$ and e_t.

Given the law of motion for A_t, one can directly solve this system for $\boldsymbol{x}_t$ as a function of the past and current shocks $\{...,\epsilon_{t-2},\epsilon_{t-1},\epsilon_t\}$. However, one can generally not solve for κ and ξ in closed form, so such an approach delivers limited insights. Instead, I provide insight in several steps using simple examples.

The first step is to combine the second and third equation in (13), to get:

$$\boldsymbol{x}_t = \underbrace{(1-\kappa)\boldsymbol{e}_t}_{\text{Bank capital channel}} - \underbrace{\xi E_t[a_{t+1}]}_{\text{Expected productivity channel}}. \tag{14}$$

This nicely identifies the two main forces, or channels, at play in the model. As we will see, the first term captures how productivity affects $\boldsymbol{x}_t$ through its effect on bank capital, both directly (i.e. because e_t is the numerator in $x_t = e_t/b^*_t$) and indirectly (e_t also affects b^*_t, with elasticity κ). I dub this channel the aggregate *bank capital channel*. The second term captures how productivity affects $\boldsymbol{x}_t$ through its effect on expected productivity (which affects b^*_t with elasticity ξ). I dub this channel the *expected productivity channel*.

The expected productivity channel Let me start with the latter, which is more straightforward. First, recall that $\xi > 0$ (Lemma 1), so that the expected productivity channel is negative. Second, consider a single shock $\epsilon_0 = 1$, so that, starting from steady state, $a_0 = 1$, and $a_t = \rho^t$, for all $t > 0$. Then, the expected productivity channel corresponds to a negative and decaying impulse response functions:

$$-\xi E_t[a_{t+1}] = -\xi\rho^t.$$

This is intuitive: the higher expected productivity, the higher the efficient level of lending (b^*). Keeping e_t constant, this means that a positive productivity shock calls for an immediate loosening of the capital requirement, and then a gradual reversion to the steady state level.

The aggregate bank capital channel One way to neutralize the expected productivity channel is to consider *iid* shocks (i.e., $\rho = 0$). Then, reproducing the exercise above, we directly get an impulse response function for the capital requirement:

$$\boldsymbol{x}_t = (1-\kappa)(\alpha\kappa)^t.$$

The shock has a direct, positive impact on e_0. This affects x_0 directly, with elasticity 1, and indirectly (through b^*_0), with elasticity κ. So, $\boldsymbol{x}_0 = 1-\kappa$. Since $\kappa < 1$ (Proposition 3), the optimal response to the shock is to immediately tighten the capital requirement. The capital requirement is tightened, but b^* is still increasing in e, so lending expands (with elasticity κ). More lending today means more physical capital and thus higher wages tomorrow (with elasticity α). Young bankers will therefore be wealthier (by $\alpha\kappa$%), hence $\boldsymbol{e}_1 = \alpha\kappa$. The same logic repeats itself, and $e_t = (\alpha\kappa)^t$, which generates a decaying pattern.

Remark. When ϵ_t is *iid*, such shocks can also be interpreted as a financial shocks. This is in the sense that they affect the wealth that flows into aggregate bank capital, which impacts banking sector's intermediation capacity, without having other effects on k^*_{t+1}. From that point of view, the bank capital channel captures some notion of a financial cycle, since *iid* shocks to bank capital have endogenously persistent effects.

5.2 Which channel dominates?

The argument above suggests that if productivity shocks are not persistent, the bank capital channel will dominate, and the optimal capital requirement will be higher in good times (i.e., when productivity is high). Does the opposite hold true if shocks are very persistent? The answer is no; the reason is that the shocks have a more persistent effect on e than on b^*.

To show this, let me now assume away default costs at both the firm level ($\Delta = 0$) and the bank level ($\gamma = 0$). This is particularly useful because the social surplus function is then isoelastic in expected productivity. Hence the closed-form solution for ξ:

$$\xi = \frac{1}{1-\alpha}.$$

This assumption also implies that $\kappa = 0$. So considering again a single unit shock to a system initially at steady states, we get:

$$\boldsymbol{x}_0 = 1 - \frac{\rho}{1-\alpha}.$$

This means that $\boldsymbol{x}_0 < 0$, unless $\rho < 1-\alpha$. Hence, unless the shock has relatively low persistence, the immediate optimal response to a positive shock is to *loosen* the requirement.

However, for all $t > 1$, we have:

$$\boldsymbol{x}_t = \underbrace{a_t + \alpha\xi E_{t-1}[a_t]}_{\text{bank capital channel}} - \underbrace{\xi E_t[a_{t+1}]}_{\text{exp. productivity ch.}}.$$

The first term on the right-hand side captures the direct effect of persistent productivity on wages and, therefore, on bank capital. The second term captures a second, indirect effect: Yesterday's increase in expected productivity raised lending with elasticity ξ. This raises k today, which raises wages and bank capital with elasticity α. The third term is the productivity channel. Substituting for productivity yields:

$$\boldsymbol{x}_t = \underbrace{\frac{\rho^t}{1-\alpha}}_{\text{bank capital channel}} - \underbrace{\frac{\rho^{t+1}}{1-\alpha}}_{\text{exp. productivity ch.}} > 0.$$

Since $\rho < 1$, we can see that persistence has a stronger effect on the bank capital channel than on the expected productivity channel. This is why the former dominates at all lags in this example.

Proposition 4. *Assume default is costless ($\gamma = 0$ and $\Delta = 0$). The aggregate bank capital channel dominates the expected productivity at any lag. If the shock is not too persistent, this is also true in the period in which the shock occurs. Formally, $\forall \rho$, following a date-0 positive productivity shock, $E_0[\boldsymbol{x}_t] > 0$, $\forall t > 0$, and $\boldsymbol{x}_0 \gtreqless 0 \Leftrightarrow \rho \gtreqless 1 - \alpha$.*

Corollary 3. *When either shocks are iid, or default is costless ($\gamma = 0$ and $\Delta = 0$), positive productivity shocks generates a persistent tightening of the optimal capital requirement.*

6 Global solution and quantitative assessment

The analytical results above suggest that the optimal capital requirement is tighter during booms than in recession. However, combining default costs and persistence could still potentially overturn the result. Furthermore, the model presents non-linearities, and elasticities κ and ξ may take significantly different values as the economy gets further from the steady state. Finally, since the dynamics of bank capital play such an important role in the results, it is desirable to allow individual banks to accumulate capital over time. In this section I add such a feature to the model, and I turn to numerical methods to solve it.

First, I solve the model numerically to generate impulse response functions for x_t^*, e_t, and b_t^*. Second, I simulate the economy under x_t^* to generate booms and busts. Third, I compare x_t^* to a suboptimal policy that neglects the bank capital channel. Fourth, I add to the model a reduced-form version of Basel II's risk-weights and, in the spirit of Basel III, I use my framework to compute back-of-the-envelope estimates for optimal time-varying macroprudential capital buffers.

6.1 Methodology

Additional ingredient In practice, retained profits are an important source of capital for banks. The two-period overlapping generation structure of the baseline model rules out profit retention. This is a useful assumption because it allows for transparent analytical results. But since studying the dynamics of x^* in the general case already requires using numerical methods, I enrich the model in order to allow for such a source of bank capital.

To do so, I use a common modeling trick and impose an exogenous exit rate for banks (see the discussion in Suarez, 2010). Specifically, I assume that bankers may live more than one period: They face a constant probability of dying λ, which yields the following condition:

$$e_{t+1} \leq \eta w_{t+1} + (1-\lambda) \underbrace{[b_t r_{t+1} + e_t]^+}_{\text{realized net worth}} . \tag{15}$$

The case where $\lambda = 1$ corresponds to the baseline model.

I maintain the assumption that aggregate bank capital is scarce. This is the case in all the simulations I present in this section. As a result, under the optimal capital requirement, surviving bankers always find it optimal to invest all their wealth in bank equity. This means that condition (15) holds with equality and defines the law of motion for aggregate bank capital.

Calibration strategy I interpret one period in the model as 1 year. I choose standard values for the basic technology parameters: $\alpha = 0.35$ and $\delta = 0.05$. The choice of a value for Δ, the extra depreciation of capital due to default, is less obvious, as previous studies have used a relatively wide range of values. As a baseline, I follow Repullo and Suarez (2013) and Martinez-Miera and Suarez (2014) and set $\Delta = 0.40$. Their rationale for such a value is that it generates realistic loan losses given default (under the Basel II standardized approach, the loss given default for non-rated corporate exposure is 45%).[14]

The exogenous random variable is A_t, which I express here in terms of loan default rate: $D_t \equiv 1 - A_t$. The relevant parameters for its distribution are the constant $\overline{A}$, and the standard deviation and persistence of the shocks, σ and ρ. I calibrate these values to match the average (3.9%), the standard deviation (1.9%), and the (annual) auto-correlation (0.83) of the US bank loan delinquency rate from 1985 to 2016.[15]

This leaves me with three free parameters: λ, η, and γ. The first two drive aggregate bank capital accumulation. Parameter λ can naturally be linked to the shareholder payout ratio. I choose $\lambda = 0.064$, which is the value induced by a payout ratio of 60% for financial firms (Floyd et al. 2015) if the return on equity is 12%.[16] Given this value for λ, I pick the value of η to target a bank capital ratio of 4% in the risky steady state, which is in line with average leverage ratios for banks in the decade before the global financial crisis (Adrian and Shin, 2010). Note that given a typical average risk weight below 50%, this satisfies the Basel II risked weighted capital ratio of 8%. The last parameter (γ) is specific to my model and captures the intensity of bank default costs. In the absence of empirical estimates, I consider several values, ranging from 0 to 4. I use $\gamma = 2$ as a baseline, which yields a simulated deadweight loss of 5% of steady state GDP, on average, when the representative bank defaults. Table 1 summarizes my calibration strategy.

[14]In my model, the loss given default on a loan (in percent) is equal to $\delta + \Delta$, hence the value for Δ. The main alternative approach is to target estimates of bankruptcy costs. For instance, Carlstrom and Fuerst (1997) argue that a reasonable range for estimates that include both the direct and indirect costs is [0.2 to 0.36]. They use a value of 0.25. See Subsection 6.4 for a discussion on how this choice and that of other parameter values affect the results.

[15]Specifically, I use the end of Q4 value of the Delinquency Rate on All Loans, Top 100 Banks Ranked by Assets [DRALT100N], retrieved from FRED, the Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/DRALT100N, April 3, 2017. I calibrate the parameters to match the moments on a simulated series of 1,000,000 shocks.

[16]Formally, $\lambda = (\text{Payout Ratio}) \times \text{ROE}/(1 + \text{ROE})$. Note that studies adopting a comparable modeling approach typically use higher values for the corresponding parameter, which they link to average survival time for firms or banks (e.g, the exit rate is 0.1 in Gertler and Kiyotaki (2010) and 0.11 in Bernanke et al. (1999), and 0.2 in Martinez-Miera and Suarez (2014)).

Table 1: Calibration

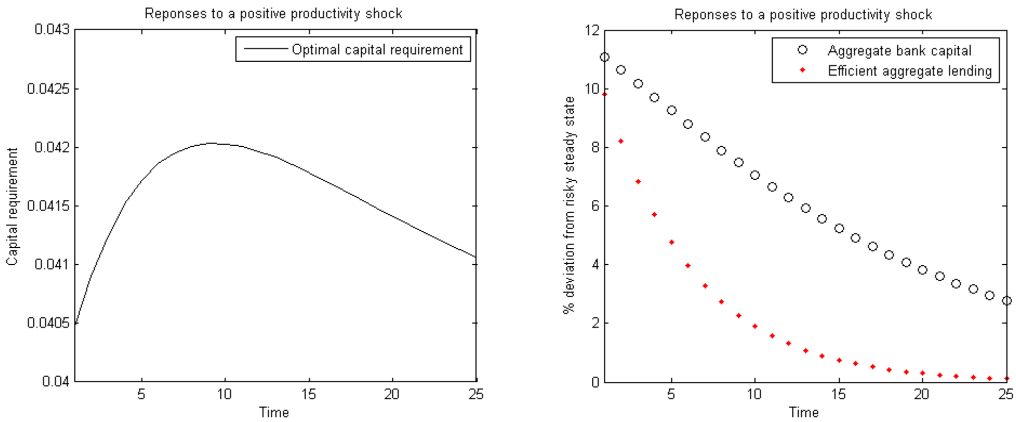
		Value	Target
Concavity of the production function	α	0.35	Capital share of output
Baseline depreciation	δ	0.05	Physical capital depreciation
Extra-depreciation in default	Δ	0.40	Loss given default on bank loans
Default rate parameters			**Loan delinquency rate (US)**
Expected success rate	$\overline{A}$	0.9575	Average (1985-2016): 3.9%
Standard deviation of the shock ϵ	σ	0.26	St. dev. (1985-2016): 1.9%
Persistence of the shock ϵ	ρ	0.83	Autocorr. (1985-2016): 0.83
Banker exit rate	λ	0.064	60% payout rate on 12% ROE
Measure of bankers	η	0.0212	4% capital ratio (8% of RWA)
Bank default cost	γ	2	I also consider $\gamma = \{0, \frac{1}{2}, 4\}$

Note: The default rate parameters are such that the moments of a simulated series of 1,000,000 shocks match those of the loan delinquency rates.

6.2 The cyclical properties of the optimal requirement revisited

Figure 1 shows the main impulse response functions for the baseline calibration. Starting from steady state, it depicts the effects of a positive productivity shock that decreases the loan default rate by 1% (from 4.25% to 3.25%). The left panel displays x_t^*, the response of the optimal capital requirement. It is initially positive (by around 5bps) and reaches 20bps after 8 periods before decaying. Since, in practice, capital requirements apply to risk-weight assets (with a 50% in my calibration), this should be interpreted as a 40bp increase in the headline capital requirement.

Figure 1: Impulse responses



This figure presents results for the baseline calibration (see Table 1). The left panel shows the path for the optimal capital requirement x_t^*, starting from steady state and following a 1% decrease in the loan default rate. The right panel decomposes the response into two parts by showing the impulse responses for aggregate bank capital (black circles) and the efficient level of aggregate lending (red dots), expressed in percent deviation from their steady state values.

The right panel shows the responses of aggregate bank capital and of the efficient level of aggregate lending. These responses are expressed in percentage deviation from steady state. They are therefore the non-linear counterparts of $\boldsymbol{e}_t$ and $\boldsymbol{b}_t$. Acknowledging a slight abuse of

Table 2: Simulation results (baseline calibration)

Mean	x_t^*	$x_t^{'}$
Unconditional	4%	3.9%
In good time	4.4%	3.2%
In bad times	2.9%	5.4%

This table compares simulated moments under the optimal capital requirement (x^*) and the alternative policy ($x^{'}$) defined by Equation (16). Both simulations use baseline calibration parameters (Table 1), start at steady state, and follow the exact same sequence of 10,000 random shocks. Good times (bad times) are defined as periods where the loan default rate is at least one standard deviation below (above) the sample mean.

notation, I will use the same bold letters to refer to these response functions. To link the two panels, note that $x_t^* - x_{ss} \simeq (\boldsymbol{e}_t - \boldsymbol{b}_t)x_{ss}$, where $x_{ss} = 0.04$. Both responses are positive and decay over time. Note that $\boldsymbol{e}_1$ is only slightly larger than $\boldsymbol{b}_1$, but the latter decays at a faster pace, which explains the hump shape of x_t^*.

The key takeaway from Figure 1 is that it confirms the key analytical insight derived above: Positive productivity shocks generates a persistent tightening of the optimal capital requirement.

Good times versus bad times Figure 1 focuses on fluctuations close to steady state. A simple way to confirm that the main insights hold globally is to simulate a full series of shocks and compare the simulated optimal requirement in good times versus bad. I define good (bad) times as periods in which the loan default rate is at least one standard deviation below (above) its mean. Table 2 displays the simulation results. As we can see, x_t^* is indeed tighter in good times (4.4%) than in bad (2.9%). These correspond to 8.8% and 5.8% capital requirements applied to risk-weighted assets (where the average risk weights are 50%).

An alternative policy To get a sense of the role of the bank capital channel for the results above, it is useful to define an alternative, suboptimal policy for the the capital requirement. I want to capture a policy that takes the expected productivity channel into account but ignores the bank capital channel. I model such a policy as follows:

$$x_t^{'} \equiv \frac{e_{ss}}{b_t^{'}}, \tag{16}$$

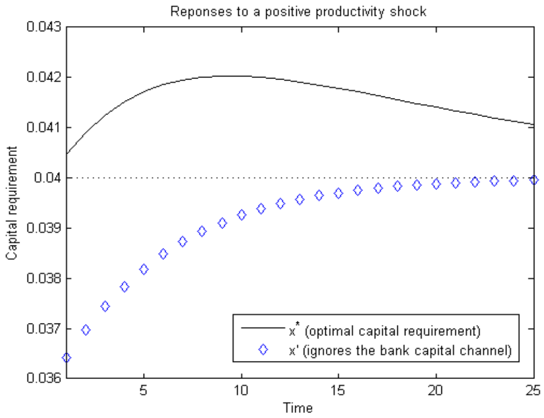
where

$$b_t^{'} \equiv \arg\max_{b_t} E_t \left[S(A_{t+1}, b_t, e_{ss}) \right].$$

That is, $b_t^{'}$ corresponds to the efficient level of aggregate lending if $e_t = e_{ss}$. This policy ignores the bank capital channel in the sense that it holds e_t constant at its steady-state level. The linearized model provides a related interpretation: Around the steady state $\boldsymbol{b}_t^{'} \simeq \xi a_t$, and therefore,

$$\boldsymbol{x}_t^{'} \simeq \boldsymbol{x}_t - \boldsymbol{e}_t(1-\kappa).$$

Figure 2: Alternative capital requirement policy



This figure presents results for the baseline calibration (see Table 1). The solid black line corresponds the the left panel of Figure 1. This is the path for the optimal capital requirement x_t^*, starting from steady state and following a 1% decrease in the loan default rate. The sequence of blue diamonds depicts the path of the capital requirement policy that ignores the bank capital channel (see equation 16).

I compare the response to a positive productivity shock under this policy to the optimal response in Figure 2. Ignoring the bank capital channel leads to a loosening of the capital requirement instead of a tightening. The difference is initially around 40bps and then decays over time. Conversely, in the case of a negative productivity shock, this policy results in excessively stringent capital requirements.

The last column of Table 2 shows that the same insights hold away from steady state: $x_t^{'}$ is looser in good times and tighter in bad times. As we will see, policy $x_t^{'}$ unnecessarily magnifies business cycle fluctuations, which resonates with the critiques of the Basel II regulatory regime. Before exploring this further, let me discuss the quantitative role of bank default costs, and present further details on the simulation results.

Bank default costs matter in bad times Table 3 compares the realized frequencies of large losses and the average probability of default for several values of γ, under the optimal policy. As we can see, going from $\gamma = 0$ to $\gamma = 4$ decreases the probability of default from 0.64% to 0.08%. The frequencies of large losses (i.e., losses larger than 10% of the bank's equity) decrease as well, but not dramatically.

Varying γ in the range considered does not have a large effect on impulse responses triggered by a single and relatively small positive shock (like those in Figure 1). Considering successive negative shocks instead clarifies the role to γ. Figure 3 illustrates this: The left panel displays the response of x_t^* to three successive negative shocks (raising the loan default rate from 4.25% to 7.24%). We can see that the optimal cut in capital requirement is less aggressive when γ is higher. On the contrary, there is no discernible difference in the response to an equivalent series of three positive shocks (see the right panel).

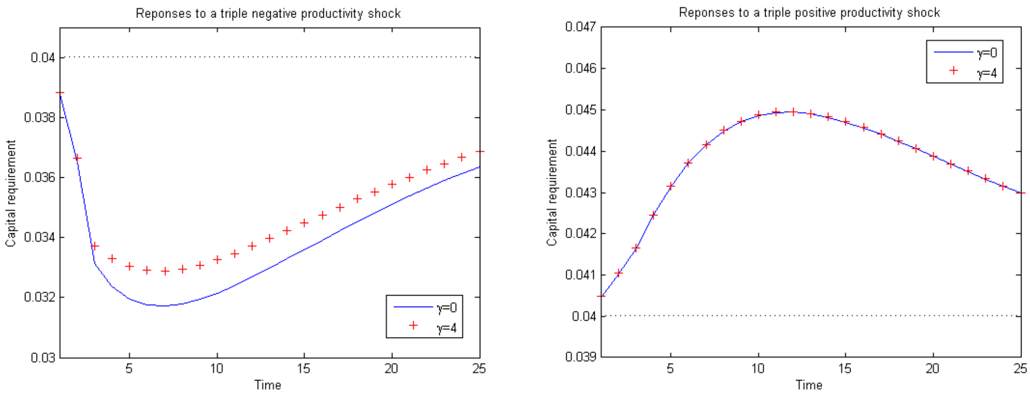
Finally, one could wonder whether the specification for the default cost function strongly affects the impulse response functions. As I show in Appendix B, this is not the case: All the

Table 3: Simulation results (baseline calibration)

	$\gamma = 0$	$\gamma = 0.5$	$\gamma = 2$	$\gamma = 4$
Freq. of large losses	17.5%	17.0%	16.4%	15.9%
in good time	7.5%	7.4%	7.1%	6.9%
in bad times	34.6%	31.8%	29.1%	26.9%
Probability of default	0.64%	0.29%	0.13%	0.08%

This table presents simulated moments for different values of γ. Other parameters are at their baseline values (Table 1). All simulations start at steady state, and follow the exact same sequence of 10,000 random shocks. Good times (bad times) are defined as periods where the loan default rate is at least one standard deviation below (above) the sample mean. Large losses correspond to those larger than 10% of initial bank capital.

Figure 3: Larger firm default costs



This figure compares the response of the optimal capital requirement to three successive negative productivity shocks (left panel) and three negative ones (right panel), for $\gamma = 0$ (solid blue line) and $\gamma = 4$ (red crosses).

insights derived in this section also apply to the case where bank default costs are proportional to the bank's balance sheet–that is, under the specification $\Psi_{t+1} = \gamma b_t$.

Comparing policies: further simulation results Table 5 in Appendix C displays a series of additional results for the comparison between x_t^* and x_t'. Here are the main findings. First, I find that aggregate lending is more volatile under x_t'. Interestingly, the frequency of large bank losses is slightly lower under this suboptimal policy (16% compared to 16.4%). However, distinguishing between good times and bad helps paint a nuanced picture. In good times, large losses are more frequent under x_t' (12.1% compared to 7.1% under x_t^*), which is consistent with the popular notion that banks are piling up risks in good times.[17] In bad times, however, banks do not take enough risk under x_t', as large losses are much less frequent than under x_t^* (11.8% compared to 29.12%). This is because x_t^* optimally trades off the costs of risk taking with the benefits of allowing for positive net present value loans to be made. This also explains why the average probability of default is slightly higher under x_t^* .

The excessive risk taking in good times under x_t' is materialized in two ways. First, banks do not fully internalize the social costs of their credit expansion, which means that they lend

[17] See, for instance, Borio and Drehmann (2009) and Boissay et al. (2016).

too much. Second, they do not fully retain their equity.[18] This feature resonates with the well-documented dramatic increase in shareholder payouts by banks in the run-up to the global financial crisis (see, e.g., Acharya et al. (2011) and Floyd et al. (2015)).

In the model, the relevant measure of welfare is economic surplus (as defined in Equation (5)). At steady state, it corresponds to aggregate consumption. I find that the average welfare loss associated with $x_t^{\prime}$ amounts to 0.04% of steady state consumption. This number is not far from those typically associated with the welfare gains from traditional stabilization policy (see e.g. Lucas (2003)).

6.3 Basel II and the countercyclical buffers of Basel III

The model is useful to think about the so-called procyclical effects of Basel II and countercyclical buffers of Basel III.

A Basel II interpretation A key feature of the Basel II regulatory regime is that bank assets are weighted for capital regulation purposes. The weights on individual loans depend on their probability of default, either directly (in the "Advanced" or "Internal Rating Based" approach) or through credit ratings (in the "Standardized" approach).

These weights were designed to deal with cross-sectional differences in loan riskiness. However, because probabilities of default co-move over the business cycle, the average, or *effective*, risk weight of a given loan portfolio evolves over time. Namely, the effective risk weight goes up during a downturn, and down during a boom. In the context of my model, a simplified version of Basel II's capital requirement can be captured by the following constraint:

$$e_t \geq \overline{x}\omega_t b_t, \tag{17}$$

where $\overline{x}$ is a constant capital requirement (for instance 8% as in the Basel II regime), ω_t is the effective risk weight on the loan portfolio at date t, and as before, e_t and b_t are the representative bank's capital and lending level. In a more general set-up, the effective risk weight would be a function of both the loan-portfolio composition (i.e., how much is lent to each firm i), and the associated probabilities of default (a_i's) and losses given default (which I could capture by indexing Δ_i's). However, since only aggregate risk matters in my model, the only relevant variable here is the expected loan default rate. Accordingly, I define

$$\omega_t \equiv \omega(E_t[D_{t+1}]),$$

[18] Under $x_t^{\prime}$, condition (15) may not be binding in equilibrium. In that case, banks pay extra dividends. Here, I allow surviving bankers to re-inject equity into the bank in the future, which somewhat alleviates the credit crunch in bad times. In reality, however, banks are typically reluctant to raise capital in bad times–because of stigma, for example, or debt overhang considerations. Such ingredients involve considerable technical challenges in a dynamic general equilibrium model (see Bahaj and Malherbe (2018)). I therefore leave this to future research.

where $\omega(\, . \,)$ is an increasing function.

Macroprudential buffers Now, consider a macroprudential regulator that takes as given the rules set by a microprudential regulator (i.e., $\overline{x}$ and ω_t). Here, the capital requirement corresponds to constraint (17), except that the macroprudential regulator can set a time-varying adjustment factor z_t. That is:

$$e_t \geq z_t \overline{x} \omega_t b_t.$$

Assuming it binds in equilibrium (which will imply $b_t = b_t^*$), efficiency requires:

$$z_t^* = \frac{e_t}{\overline{x} \omega_t b_t^*}. \tag{18}$$

Log-linearizing around the steady state gives:

$$\boldsymbol{z}_t = \boldsymbol{e}_t - \boldsymbol{b}_t - \boldsymbol{\omega}_t,$$

and the following expression for the optimal macroprudential buffer:

$$z_t^* \simeq \boldsymbol{z}_t \overline{x}.$$

Since ω increases with the loan default rate, the response of $\boldsymbol{\omega}_t$ to a positive productivity shock is negative. Hence, ceteris paribus, this loosens the capital constraint. From a conceptual point of view this is not necessarily a bad thing, because higher productivity calls for more investment, which requires more lending. In the context of the model, such a desired increase in lending due to the increase in average credit quality (a lower loan default rate) can be captured by $\boldsymbol{b}_t^{'}$, which leads to the following interpretation:

$$\boldsymbol{z}_t \simeq \underbrace{\boldsymbol{e}_t(1-\kappa)}_{\text{bank capital}} - \underbrace{\boldsymbol{b}_t^{'}}_{\text{credit quality}} - \underbrace{\boldsymbol{\omega}_t}_{\text{risk weights}}.$$

It is widely accepted that the time-varying effect of risk weights is excessive. There is no empirical consensus on an estimate for the elasticity of ω_t with respect to productivity, but a series of studies have attempted to measure such an effect. For instance, using a default probability model, Kashyap and Stein (2004) obtain an implied increase of 10% in risk weights from 1998 to 2002 (a period in which the US economy contracted and in which, according to the data I use for calibration, there was an increase of a bit less than one percent in delinquency rates).[19] Accordingly, I will assume $\boldsymbol{\omega}_4 = -10\%$.

When changes in γ have little effect on $\boldsymbol{x}_t$ this means that $\kappa \simeq 0$. From the discussion above, we now that this is a good approximation in normal and good times. The baseline im-

[19]Other approaches typically give higher numbers. Kashyap and Stein (2004) propose several approaches and discuss other estimates found in the literature (see, e.g., Catarineu-Rabell et al. (2005)).

pulse response (Figure 1) can then be used in a back-of-the-envelope estimation of the optimal macroprudential buffer. So, I set $\boldsymbol{e}_4 = 9.7\%$ and $\boldsymbol{b}_4 = 5.7\%$.

I can then compute:

$$z_4 \simeq \underbrace{9.7\%}_{\text{bank capital}} - \underbrace{5.7\%}_{\text{credit quality}} - \underbrace{-10\%}_{\text{risk weights}} = 14\%.$$

Multiplying by $\overline{x} = 8\%$ gives the optimal macroprudential buffer:

$$z_4^* = \underbrace{78\text{bps}}_{\text{bank capital}} - \underbrace{46\text{bps}}_{\text{credit quality}} + \underbrace{80\text{bps}}_{\text{risk weights}} = 112\text{bps}$$

Thus, I find that a 1% decrease in the default rate for firms should be met, at the 4-year horizon, by an increase of the requirement from 8% to 9.12%. This can be interpreted as a 112bps countercyclical capital buffer. The 112bps correspond to 80bps needed to neutralize the decrease in risk-weights, plus 78bps to account for the bank capital channel, minus 56bps to account for the improved credit quality (which reflects the expected productivity channel). This suggests that accounting for the bank capital channel may be quantitatively as important as accounting for the procyclical effects of risk weights. Additionally, this suggests that the welfare gains from financial stabilization policy could be twice as large as those I find in the exercise.

6.4 Sensitivity analysis

I have explored the parameter space beyond the set of values of my baseline calibration. The general conclusion from my numerical analysis is that the shape of the impulse response functions and, to some extent, their magnitude are quite robust to alternative parametrizations.

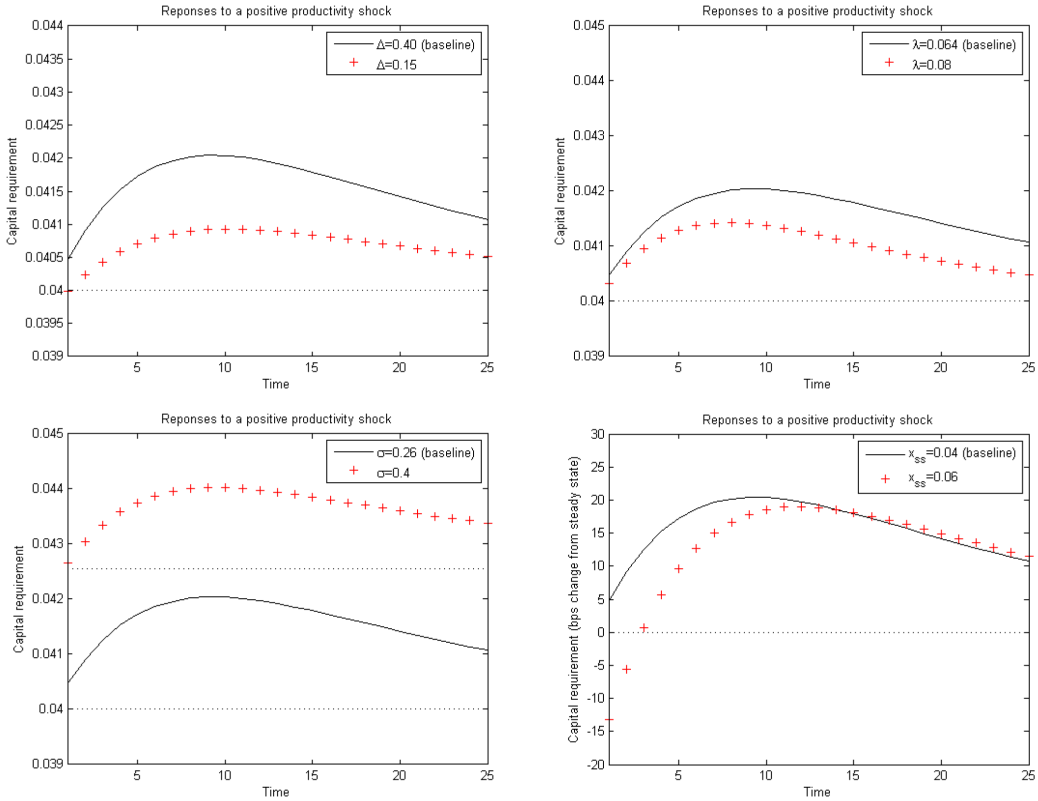
Figure 4 illustrates how changes in the value of non-standard parameters values affect the impulse response for x_t^*. In each case, we can see that the change tends to dampen the response function, but it remains positive and hump shaped. The top-left panel considers $\Delta = 0.15$, which is much lower than the baseline ($\Delta = 0.4$) and in the lower part of the range used in the literature.

The top-right panel considers a higher value for λ (0.08 instead of 0.064). According to my shareholder payout approach, this could either correspond to a steady state ROE of 15.5% (keeping the shareholder payout ratio constant at 60%) or to a steady-state shareholder payout ratio of 75% (keeping the ROE constant at 12%). Raising the value of λ further continues to dampen the response, but brings about steady-state values for the shareholder payout ratio and the ROE that are arguably not reasonable.[20]

The bottom-left panel considers an increase in σ (the standard deviation of the shock). This

[20] At values above 0.25, the response is small (a couple of basis points) and starts in negative territory.

Figure 4: Parameter sensitivity



This figure illustrates how the impulse response for x_t^* is affected by a change of value for parameters Δ, λ, σ, and η. In all panels but the bottom right, the solid line depicts x_t^* in the baseline calibration (as in Figure 1), and the sequences of red crosses depicts x_t^* for an alternative value of the considered parameter. The bottom-right panel compares the impulse response for two different steady-state capital ratio targets: 4% (baseline, solid line) and 6% (alternative, red crosses). The alternative target is achieved by adjusting parameter η. The responses are normalized at their corresponding steady state level.

has two effects. The first one is to raise the steady-state optimal requirement; the second is to slightly dampen the response.

Finally, the bottom-right panel displays the effect of a change in the target for the steady state capital ratio (it keeps λ constant and lets η adjust to meet the target). This substantially affects the initial response as it becomes negative. However, the response quickly becomes positive and converges to the baseline case. This pattern is very similar to the one described in Proposition 4.

7 Discussion

7.1 Efficiency benchmark and regulatory objective

Since the economy has heterogeneous agents (savers and bankers, of different generations), the choice of an efficiency benchmark is not trivial.

Bank and firm default costs are meant to capture the deadweight losses associated with agency problems. One can envision the first best level of investment as the one that would maximize economic surplus if these costs did not exist (i.e., $\gamma = \Delta = 0$). In the presence of costs, however, such a level is not an appropriate benchmark. In contrast, the benchmark I use takes default costs into account. It can thus be interpreted as a second best. The market failure that I establish is that the competitive equilibrium is worse than the second best (and exhibits a higher investment level).

In a constrained efficiency problem, the regulator faces the same constraints as the private agents, and aims at implementing a second-best outcome, which must be Pareto efficient (without arbitrary transfers). This is the approach I follow. If I allowed for transfers, the regulator would make sure that, ex ante, bankers would have enough wealth to fully alleviate deadweight losses. To finance this, it would initially tax savers. Then, ex post, it would tax bankers to finance a compensating transfer the savers.[21] This would make the problem less interesting, but the main results would still hold, as they hold for $\gamma = 0$.

For computational convenience, I assume that bailout taxes are imposed on savers. If taxes are instead imposed on workers, then the allocation attained in the constrained optimum is an ex-ante Pareto improvement compared to the competitive equilibrium (and all the main insights go through). This makes the regulatory objective particularly appealing, as economic surplus maximization does not require that any agent ends up strictly worse off. Note also that given the regulatory objective, Proposition 2 establishes that focusing on capital requirements is not

[21] The same logic implies that having the regulator maximize an arbitrary intertemporal welfare function could be problematic. A simple example illustrates why. Imagine that the regulator would like to transfer wealth from current generation to the next period's banker. He may then be tempted to allow for over-investment today, as a larger capital stock will boost tomorrow's wage and, therefore, the wealth of future bankers. This involves negative net-present-value investment, but the regulator could still find it desirable given its objective functions and the constraints it faces.

restrictive, since they allow to achieve efficiency at all dates.

Finally, note that the friction at the firm level (costly verification), contributes to the fact the constrained efficient level (k^*) is strictly lower than the first best. However, this friction does not generate a market failure (the debt contract is an optimal contract and verification costs are not affected by general equilibrium effects). To see that, consider a version of the model without deposit insurance or bank default costs. It is straightforward to show that the competitive equilibrium will correspond to the planner's solution (both allocations will be pinned down by $\alpha E[A]k^{\alpha-1} = 1$).

Additional frictions at the firm level could create a market failure. If the frictions point in the direction of inefficient over-investment, this would not affect much the analysis. However, if they point in the direction of under-investment and if they more than offset the effect of the other frictions, then the problem of the regulator becomes to stimulate investment. In this case, capital requirements will generally not be enough to restore investment efficiency (more on this below).

7.2 Default cost functions: properties and interpretation

There are several possible interpretations for my specification of Ψ. As mentioned in Section 4, one can think of deadweight losses as stemming from the bank resolution process.[22] Arguably, this process will be longer and costlier when financial distress is high. Alternatively, one can think of default costs as a proxy for the deadweight losses of taxation associated with bailouts (see, e.g., Acharya et al. (2010) and Bianchi (2013)).

In my specification, realized costs are linear in financial distress and, therefore, in the amount needed for a bailout (this is also the assumption in Acharya et al. (2010)). Hence, it is not the case that financial distress at a bank increases, in itself, the likelihood or the extent of other banks' financial distress. It is therefore quite different from a fire-sale externality. That said, making γ an increasing function of aggregate financial distress would be an easy way to add a fire-sale flavor to the model.

Finally, note that my specification corresponds to the baseline setup in which Townsend (1979) establishes the optimality of standard debt contracts in a costly-state-verification environment.[23]

Alternative specification Ψ_2: balance sheet size Townsend (1979) also considers fixed

[22]Bankruptcies often involve deadweight losses (Townsend (1979)). In the case of financial institutions, losses can be large (James (1991)), and banking crises are typically followed by long and painful recessions (Reinhart and Rogoff (2009)) involving permanent output losses (Cerra and Saxena (2008)).

[23]Specifically, the equivalence is as follows. In Townsend (1979) and following his own notation, the assumption for his Proposition 3.1 is that verification costs are increasing and convex in $\bar{C} - \bar{g}(y_2)$, where $\bar{C}$ is the repayment in absence of default (i.e., when no verification occurs), and $\bar{g}(y_2)$ is the repayment in the case of default (i.e. when verification occurs). Townsend then goes on with the example where verification costs are simply $\lambda(\bar{C} - \bar{g}(y_2))$, where λ is a positive constant, which is thus equivalent to my specification for Ψ.

costs. The problem he studies is of fixed size. In the present context, the natural interpretation of a fixed verification cost is one that is proportional to the size of the bank's balance sheet. Formally, conditional on default, this gives:

$$\Psi_2 \equiv \begin{cases} \gamma b & ; v < 0 \\ 0 & ;\text{ otherwise} \end{cases}$$

In this case, realized default costs are not affected by lending decisions at other banks (Ψ_2 does not depend on k). However, probabilities of default are affected (A_0 still depends k). This has two implications:

First, expected default costs also exhibit decreasing return to scale in the aggregate. However, in contrast to Ψ , Ψ_2 is discontinuous at the default threshold. This makes the problem less analytically well behaved. As a result, one can no longer derive a result as strong as Proposition 3, as one cannot rule out local jumps in the $b^*(e)$ function.[24] Still, as I have explained in the previous section, using specification Ψ_2 in the numerical analysis does not substantially affect the results.

Second, under Ψ_2, it is also true that the competitive equilibrium exhibits over-lending, even in the absence of government guarantees. The relevant externality is simple and very intuitive: Banks do not internalize that by extending lending they increase the probability of default of other banks. In my specification, the same is true, but this does not create a wedge in the first order condition (because $\Psi(A_0) = 0$). What creates the wedge is the externality that works through realized default costs.[25]

Alternative specification Ψ_3: value of distressed assets Finally, I consider the case where default costs are proportional to the value of distressed assets. That is, conditional on default:

$$\Psi_3 \equiv \gamma b(1+r)$$

Since Ψ_3 depends on k (through r), realized default costs are affected by the general equilibrium effect. Taking this into account, we have:

$$\frac{d\Psi_3}{db} = \gamma(1+r) + \underbrace{\gamma b r'(b)}_{<0}$$

[24]The same applies to $b^*(E[A])$ and, therefore, to Lemma 1. To see the problem, consider the case (arguably extreme) in which A_0 is to the right of the mode of the distribution. In this case, shifting the distribution to the right *increases* the density at A_0. This means that, starting at a given initial level, a marginal increase in b generates a *larger* increase in the probability of incurring the fixed cost. This is not sufficient per se to make b^* decreasing in $E[A]$, but it can contribute to it.

[25]In the absence of government guarantees, banks face interest rates that depend on their default probability and on the creditors' losses given default. Under my specification for Ψ and under Ψ_2 this only reinforces the negative externality.

Assume that there isn't deposit insurance, so that the bank does care about expected default costs. Still, the third term above does not appear in its first order condition (in contrast to that of the planner). It is negative, which means that bankers do not internalize that their credit expansion *decreases* the default cost for other banks because it increases the losses they suffer on their loans.

Now, imagine that the distribution for A is binary: $A \in \{A_L, A_H\}$, with $\Pr(A_L) = p$, and assume that, in the competitive equilibrium, the representative bank defaults in the bad state. Then, a small increase in b will not affect the probability of default (it will still be p). But it will decrease realized default costs. This means that lending embeds a *positive* externality and that the representative bank under-lends in the competitive equilibrium.

In this case (which, recall, assumes away deposit insurance), a capital requirement would not necessarily restore efficiency (although it could if the global maximum for the planner implies a zero default probability for banks). However, the planner can achieve it, in a Pigouvian approach, with a subsidy to lending, that just offsets the externality.[26]

Under a more general distribution function for A, the general equilibrium effect also acts through the probability of default. As we have seen, this tends to make the externality negative instead. In that case, then, one can generally determine the sign of the externality. Either under- or over-investment can ensue.

A specification equivalent to Ψ_3 was used in Bernanke and Gertler (1989) and in many papers of the *financial accelerator* literature that followed (e.g., Carlstrom and Fuerst (1997); Bernanke et al. (1999)). Even though the structure of these models is different, one can conjecture that the positive externality I identify above plays a role in making downturns persistent (which is a typical feature in that literature). The logic would go as follows. When firms operate at a low scale (perhaps because this is a condition for obtaining credit), they increase the expected default costs of other firms (because the equilibrium price of capital goes up), thereby limiting their access to credit. To the best of my knowledge, such a mechanism has not been suggested yet. If this conjecture is correct, it suggests that, even statically, models with financial accelerators could deliver constrained inefficient outcomes (the literature typically considers longer term notions of welfare).

8 Conclusion

I have studied economies where banks do not fully internalize the social costs of lending. In a competitive equilibrium, this translates into over-investment by firms. The regulator can restore investment efficiency thanks to a time-varying capital requirement. A key feature is that it is increasing in aggregate bank capital. A regulatory regime that overlooks such a bank capital channel will unnecessarily magnify business and financial cycle fluctuations. A calibration of

[26] Specifically, the subsidy rate must be: $p\gamma b r'(b)$.

the model suggests that the associated welfare losses are substantial. I therefore believe that studying the real-world determinants of aggregate bank capital accumulation is an important avenue for future research.[27]

References

Acharya, Viral V and Matthew Richardson, "Causes of the financial crisis," *Critical Review*, **2009,** *21* **(2-3), 195–210.**

__ , Deniz Anginer, and A Joseph Warburton, "The end of market discipline? investor expectations of implicit state guarantees," 2016. mimeo, NYU.

__ , Irvind Gujral, Nirupama Kulkarni, and Hyun Song Shin, "Dividends and bank capital in the financial crisis of 2007-2009," *NBER Working Paper Series*, **2011.**

__ , João AC Santos, Tanju Yorulmazer et al., "Systemic risk and deposit insurance premiums," *Economic Policy Review*, **2010, (Aug), 89–99.**

Admati, A.R., P.M. DeMarzo, M.F. Hellwig, and P. Pfleiderer, "Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity is Not Expensive," 2010.

Adrian, Tobias and Hyun Shin, "Liquidity and Leverage," *Journal of Financial Intermediation*, **July 2010,** *19* **(3).**

Allen, Franklin, Elena Carletti, and Robert Marquez, "Deposits and bank capital structure," *Journal of Financial Economics*, **2015,** *118* **(3), 601–619.**

Atkeson, Andrew G, Adrien d'Avernas, Andrea L Eisfeldt, and Pierre-Olivier Weill, "Government Guarantees and the Valuation of American Banks," in "NBER Macroeconomics Annual 2018, volume 33," University of Chicago Press, 2018.

Bahaj, Saleem and Frederic Malherbe, "The Forced Safety Effect: How Higher Capital Requirements Can Increase Bank Lending," 2018.

BCBS, "Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework," *Basel Committee on Banking Supervision*, **2004.**

__ , "Basel III: A global regulatory framework for more resilient banks and banking systems," *Basel Committee on Banking Supervision*, **2010.**

Begenau, Juliane, "Capital Requirement, Risk Choice, and Liquidity Provision in a Business Cycle Model," 2015. mimeo, HBS.

Bernanke, Ben and Mark Gertler, "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review*, **March 1989,** *79* **(1), 14–31.**

__ , __ , and Simon Gilchrist, "The financial accelerator in a quantitative business cycle framework,," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, **(Elsevier) 1999.**

[27] Our current understanding of bank dynamic capital structure decisions is at best incomplete, especially in general equilibrium (see the discussions in Allen et al. (2015), and Repullo and Suarez (2013) for instance, and Rampini and Viswanathan (2018), He and Krishnamurthy (2011), and Coimbra and Rey (2018) for recent advances).

Bianchi, Javier, "Efficient bailouts?," 2013. mimeo, University of Wisconsin.

Boissay, Frédéric, Fabrice Collard, and Frank Smets, "Booms and banking crises," *Journal of Political Economy*, 2016, *124* (2), 489–538.

Borio, C. and M. Drehmann, "Assessing the risk of banking crises–revisited," *BIS Quarterly Review*, 2009, pp. 29–46.

Carlstrom, Charles T and Timothy S Fuerst, "Agency costs, net worth, and business fluctuations: A computable general equilibrium analysis," *The American Economic Review*, 1997, pp. 893–910.

Catarineu-Rabell, Eva, Patricia Jackson, and Dimitrios P Tsomocos, "Procyclicality and the new Basel Accord-banks' choice of loan rating system," *Economic Theory*, 2005, *26* (3), 537–557.

Cerra, Valerie and Sweta Chaman Saxena, "Growth dynamics: the myth of economic recovery," *The American Economic Review*, 2008, *98* (1), 439–457.

Coeurdacier, Nicolas, Hélène Rey, and Pablo Winant, "The Risky Steady State," *The American Economic Review*, 2011, *101* (3), 398–401.

Coimbra, Nuno and Hélene Rey, "Financial Cycles with Heterogeneous Intermediaries," Working Paper 23245, NBER 2018.

Collard, Fabrice, Harris Dellas, Behzad Diba, and Olivier Loisel, "Optimal monetary and prudential policies," *American Economic Journal: Macroeconomics*, 2017, *9* (1), 40–87.

Dewatripont, Mathias and Jean Tirole, "Macroeconomic shocks and banking regulation," *Journal of Money, Credit and Banking*, 2012, *44* (s2), 237–254.

Diaz, Antonia and Luis Franjo, "Capital goods, measured TFP and growth: The case of Spain," *European Economic Review*, 2015.

Fernández-Villaverde, Jesús, Luis Garicano, and Tano Santos, "Political credit cycles: the case of the Eurozone," Technical Report 3 2013.

Floyd, Eric, Nan Li, and Douglas J Skinner, "Payout policy through the financial crisis: The growth of repurchases and the resilience of dividends," *Journal of Financial Economics*, 2015, *118* (2), 299–316.

Garcia-Santana, Manuel, Enrique Moral-Benito, Josep Pijoan-Mas, and Roberto Ramos, "Growing Like Spain: 1995-2007," Technical Report, CEPR Discussion Papers 2016.

Gersbach, Hans and Jean-Charles Rochet, "Capital regulation and credit fluctuations," *Journal of Monetary Economics*, 2017, *90*, 113–124.

Gertler, M. and N. Kiyotaki, "Financial intermediation and credit policy in business cycle analysis," *Handbook of Monetary Economics*, 2010, *3*, 547.

He, Zhiguo and Arvind Krishnamurthy, "A model of capital and crises," *The Review of Economic Studies*, 2011.

James, Christopher, "The losses realized in bank failures," *The Journal of Finance*, 1991, *46* (4), 1223–1242.

Jordà, Òscar, Moritz Schularick, and Alan M Taylor, "When credit bites back," *Journal of Money, Credit and Banking*, 2013, *45* (s2), 3–28.

Kareken, J.H. and N. Wallace, "Deposit insurance and bank regulation: A partial-equilibrium exposition," *Journal of Business*, 1978, pp. 413–438.

Kashyap, A.K. and J.C. Stein, "Cyclical implications of the Basel II capital standards," *Economic Perspectives - Federal Reserve Bank of Chicago*, 2004, *28* (1), 18–33.

Laeven, Luc and Fabian Valencia, "Systemic banking crises database: An update," 2012.

Lucas, Robert, "Macroeconomic priorities," *The American Economic Review*, 2003, *93* (1), 1.

Malherbe, Frederic and Michael McMahon, "Beyond Pangloss: Financial sector origins of inefficient economic booms," Technical Report, mimeo 2019.

Martinez-Miera, D. and J. Suarez, "A Macroeconomic Model of Endogenous Systemic Risk Taking," 2014. mimeo, CEMFI.

Morrison, Alan D and Lucy White, "Crises and capital requirements in banking," *American Economic Review*, 2005, pp. 1548–1572.

Perotti, E.C. and J. Suarez, "Last bank standing: what do I gain if you fail?," *European economic review*, 2002, *46* (9), 1599–1622.

Rampini, Adriano A and S Viswanathan, "Financial intermediary capital," *The Review of Economic Studies*, 2018, *86* (1), 413–455.

Reinhart, Carmen M and Kenneth Rogoff, *This time is different: Eight centuries of financial folly*, Princeton University Press, 2009.

Repullo, Rafael, "Cyclical adjustment of capital requirements: A simple framework," *Journal of Financial Intermediation*, 2013, *22* (4), 608–626.

_ and Javier Suarez, "The procyclical effects of bank capital regulation," *Review of Financial Studies*, 2013, *26* (2), 452–490.

_ , Jesús Saurina, and Carlos Trucharte, "Mitigating the Pro-cyclicality of Basel II," *Economic Policy*, 2010, *25* (64), 659–702.

Rognlie, Matthew, Andrei Shleifer, and Alp Simsek, "Investment hangover and the great recession," *American Economic Journal: Macroeconomics*, 2018, *10* (2), 113–53.

Santos, Tano, "Antes del diluvio: The Spanish banking system in the first decade of the euro," *manuscript, Columbia University*, 2014.

Suarez, Javier, "Discussion of: Procyclicality of capital requirements in a general equilibrium model of liquidity dependence," *International Journal of Central Banking, forthcoming*, 2010.

Townsend, Robert M, "Optimal contracts and competitive markets with costly state verification," *Journal of Economic theory*, 1979, *21* (2), 265–293.

Van den Heuvel, Skander J, "The welfare cost of bank capital requirements," *Journal of Monetary Economics*, 2008, *55* (2), 298–320.

Appendix A: Proofs

Proposition. *1. At all dates, the competitive equilibrium capital stock is inefficiently high. That is:* $k_{t+1}^{CE} > k_{t+1}^{*}, \forall t.$

Proof. The representative banker maximizes $E_t\left[[b_t r_{t+1} + e_t]^+\right] - e_t$ with respect to $e_t, b_t \geq 0$. The respective first order conditions are:

$$\begin{cases} \int_{\rho_{t+1}}^{\infty} f_t(r_{t+1}) dr_{t+1} \leq 1 \\ \int_{\rho_{t+1}}^{\infty} r_{t+1} f_t(r_{t+1}) dr_{t+1} \leq 0 \end{cases},$$

where f_t is the probability distribution function of r_{t+1} conditional to date-t information, and $\rho_{t+1} \equiv -\frac{e_t}{b_t}$ is the solvency threshold. That is, this is the value for r_{t+1} below which the bank defaults..

Efficiency requires that $E_t[r_{t+1}] \geq 0$. Given the banker's first order condition, $E_t[r_{t+1}] > 0$ cannot be true in equilibrium. The only case where the competitive equilibrium could be efficient is when efficiency requires $E_t[r_{t+1}] = 0$. Assume that this is the case and consider the first order conditions.

If $E_t[r_{t+1}] = 0$, the only way for both conditions to be satisfied is for the bank never to default. If this is the case, the bank is locally indifferent between financing the loans with deposits or equity. Now, imagine that the bank substitutes deposits for equity (keeping lending constant) up to the point where it fails if a low A_t is realized. From this point on, decreasing equity further increases expected profits (for each unit of equity withdrawn today, less than a unit is repaid to depositors tomorrow in expectation). Hence, the bank defaults with strictly positive probability in equilibrium and, given the first order conditions, $E_t[r_{t+1}] = 0$ cannot be satisfied in equilibrium. □

Proposition. *2. The regulator can ensure investment efficiency at all t with the following capital requirement:*

$$x_t^* = \frac{\eta_t w_t}{k_{t+1}^*}. \tag{19}$$

Proof. Given that aggregate bank capital is assumed to be scarce (i.e., $\eta_t w_t < k_{t+1}^*$) and that efficiency requires $E_t[r_{t+1}] \geq 0$, bankers find it profitable to invest all their wealth in equity. Assume that the representative bank defaults in equilibrium. Then, following the same logic as in the proof of Proposition 1, the banker always prefers to finance loans with deposits at the margin. Hence the capital requirement must be binding, and $k_{t+1}^* = e_t / x_t^*$ in equilibrium. If the representative bank does not default in the constrained equilibrium, a similar logic applies, but the optimal capital requirement is not unique. □

Lemma. 1. *Denote ξ the point elasticity of b^* with respect to $E[A]$. For all $E[A]$, $\xi(E[A]) > 0$.*

Proof. If $\gamma = 0$, we have $b_t^* = \left(\frac{\alpha E_t[A_{t+1}]}{\delta + (1 - E_t[A_{t+1}])\Delta}\right)^{\frac{1}{1-\alpha}}$, and the result is obvious. With $\gamma > 0$, it suffices to verify that $\Psi(E[A], b)$ is submodular. □

Proposition. *3. Denote κ the point elasticity of b with respect to e. For all $e \geq 0$, $\kappa(e) < 1$.*

Proof. Since $\Psi(A_0, b) = 0$, the first order condition (10) reads:

$$\alpha A(b)^{\alpha-1} - (\delta + (1-A)\Delta))\, b + \gamma \int_0^{A_0(b,e)} \left(r(b) + br^{'}(b) \right) f(A) dA = 0.$$

The left-hand side defines a function $\overline{S}_b(b,e)$. Since the integrand is decreasing in b, and A_0 is increasing in b and constant in leverage (e/b), this function exhibits decreasing returns to scale. Since $\overline{S}_1(b^*, e) = 0$ and $\overline{S}(b,e)$ is concave in b, it follows that:

$$\overline{S}_1(\lambda b, \lambda e) < \lambda \overline{S}_1(b,e) \Rightarrow b^*(\lambda e) < \lambda b^*(e) \Rightarrow \kappa(e) < 1.$$

□

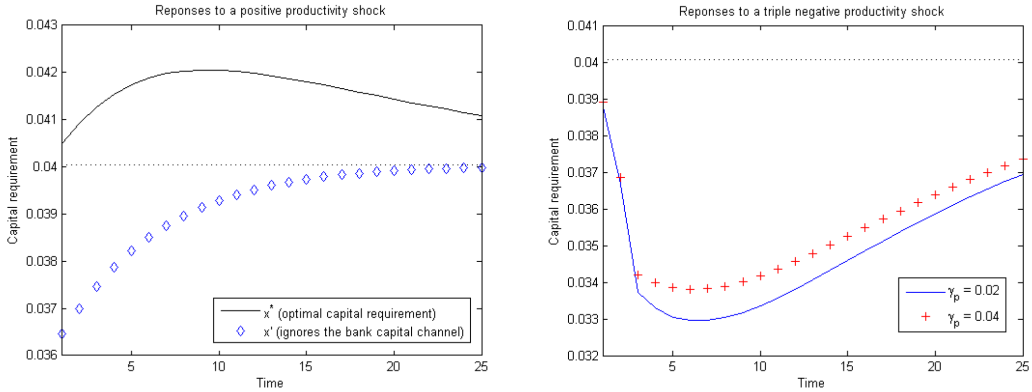
Proposition. *4. Assume default is costless ($\gamma = 0$ and $\Delta = 0$). The aggregate bank capital channel dominates the expected productivity at any positive lag. If the shock is not too persistent, this is also true in the period in which the shock occurs. Formally, following a date-0 positive productivity shock, $E_0[\boldsymbol{x}_t] > 0$, $\forall t > 0$, and $\boldsymbol{x}_0 \lesseqgtr 0 \Leftrightarrow \rho \lesseqgtr 1 - \alpha$.*

Proof. The proof follows exactly the same logic as the example in the text. The only difference is that one needs to take expectations into account for the realizations of ϵ_t for $t > 0$. □

Appendix B: Alternative default cost function

Figure 5 and Table 4 present results under the alternative bank default cost function (specification $\Psi_{t+1} = b_t \gamma_p$). The value for the cost parameter is chosen so that the expected costs given default are comparable to those under the baseline calibration. That is, $\gamma_p = 0.02$, which yields an expected deadweight loss (conditional on default) of 5.88% of steady state GDP. As we can see, both the impulse responses and the loss distributions are very similar to what I get with my baseline specification (see Figures 1 and 2 and Table 3). Finally, note that the welfare loss is essentially unchanged (it remains around 0.04% of steady state consumption).

Figure 5: Bank default costs proportional to balance sheet



This figure presents the results under the assumption that $\Psi_{t+1} = 0.02b_t$ if the bank defaults, and 0 otherwise. The left panel shows the path for the optimal capital requirement x_t^*, and the alternative policy, starting from steady state and following a 1% decrease in the loan default rate. The right panel shows the path for the optimal capital requirement after a triple negative shock for two different values of γ_p.

Table 4: Loss frequencies

	$\gamma = 2$	$\gamma = 4$	$\gamma_{\text{prop}} = 0.02$
Freq. of large losses	16.4%	15.9%	16.0%
in good time	7.1%	6.9%	7%
in bad times	29.1%	26.9%	28%
Probability of default	0.13%	0.08%	0.11%

This table compares simulated moments between the baseline specification for the default cost function (with $\gamma = 2$ and $\gamma = 4$) and the alternative specification $\Psi_{t+1} = 0.02b_t$. Other parameters are at their baseline value (Table 1). All simulations start at steady state, and follow the exact same arbitrary sequence of 10,000 random shocks. Good times (bad times) are defined as periods in which the loan default rate is at least one standard deviation below (above) the sample mean. Large losses correspond to those larger than 10% of initial bank capital.

Appendix C: Comparing policies

Table 5: Further simulation results (baseline calibration)

		x_t^*	$x_t^{'}$
Lending volatility	$\mathrm{std}(b_t)$	1.98	2.19
Probability of default	$\mathrm{mean}(\Pr_t(v_{t+1} < 0))$	0.13%	$\simeq 0$
Freq. of large losses	ROE $< -10\%$	16.4%	16.0%
	in good time	7.1%	12.1%
	in bad times	29.1%	11.8%
Dividend / recap	$(\hat{e}_t - e_t)/\hat{e}_t$	0	1.6%
	in good time	0	7.1%
	in bad times	0	−3.4%
Welfare loss	$\mathrm{mean}\left(S_t^* - S_t^{'}\right)/S_{ss}$	-	0.04%

This table compares simulated moments under the optimal capital requirement (x^*) and the alternative policy ($x^{'}$). Both simulations use baseline calibration parameters (Table 1), start at steady state, and follow the exact same arbitrary sequence of 10,000 random shocks. Good times (bad times) are defined as periods where the loan default rate is at least one standard deviation below (above) the sample mean. ROE is defined as $\left(v_{t+1}^{+}/e_t\right) - 1$. Variable $\hat{e}_t$ denotes the pre-dividend (and pre-recapitalization) level of aggregate bank equity at the end of date t. Finally, S_t^* and $S_t^{'}$ are the realized surplus under the respective policies.