Bank Capital Requirements and Bank Lending

Saleem Bahaj, Chiara Lattanzio, and Frederic Malherbe[1]

Abstract: Macroprudential regulation is often viewed as a trade-off between banking system stability and aggregate credit supply. In this paper, we provide a comprehensive analysis of how changes in capital requirements affect bank lending. We use a theoretical framework to assess and nuance the trade-off. We show that imperfect competition, general equilibrium effects, and asset heterogeneity among banks result in lending responses that are complex and difficult to estimate. Armed with these theoretical insights, we assess existing strategies in the empirical literature and provide guidance for future research.

Keywords: macroprudential regulation; capital requirements.

1 Introduction

Banks have incentives to lever up to levels that are excessive from a social point of view. Capital requirements of some variety have been used by regulators for more than a century to address, or at least contain such an issue, Haubrich (2020). Basel I, the first wave of international standards for bank capital regulation, was introduced in the 1980s. Its focus was microprudential and its main goal was to "level the playing field" across different jurisdictions. It also introduced the notion of Risk Weighted Assets, according to which the effective capital requirement on an asset depends on its risk category.

Basel I risk-weight categories were quite coarse. Basel II introduced risk weights that: i) Were based on models of credit risks; and ii) Increased with perceived riskiness. An important question quickly emerged: what are the general equilibrium effects of such regulation? Beyond its academic interest, the question reveals a policy concern: do such policies amplify the economic cycle? The basic reasoning is simple: In downturns, as provisions and losses mount up, bank capital erodes, and, at the same time, as risks rise, so do risk-weights. Hence, while the numerator of a bank's risk-weighted capital ratio goes down, its denominator goes up, and the ratio plummets. Unless the bank raises more capital (or initially had a sufficient capital buffer over and above the requirement), it will have to downsize, typically cutting credit to the economy. If the banking

[1] The authors are with University College London.

sector as a whole contracts lending, real effects can be substantial. Moreover, if banks downsize by selling assets, this could trigger a fire-sale, destabilising the financial system.

The macroprudential approach to financial regulation seeks to safeguard both the financial system as a whole and the resilience of the real economy to a financial shock. Accordingly, Basel III introduced time-varying capital requirements (the so-called countercyclical capital buffer, or CCyB) and additional buffers for large, systemically important institutions.

Better capitalized banks are safer, which improves financial stability and limits taxpayer exposure. To ensure financial stability, regulator could impose very high, even 100%, capital requirements. However, such a policy is unlikely to be optimal if it hinders credit provision. To see this, imagine that there is only one type of bank assets: loans, and there are two types of liabilities: equity (which counts as capital) and debt. Imagine, also, that the level of loans that the bank decides to make is a function of the capital requirement it faces. Say, at a 0% capital requirement, the bank makes \$100 of loans, at a 20% requirement, it makes \$90 of loans, and at a 100% requirements, only \$50. Which combination of capital requirements and loans is better for the regulator depends on how it trades off bank leverage (and financial stability) with the level of lending (and economic activity). This trade off is likely to be affected by a series of macroeconomic and financial factors and is therefore evolving over time (see, e.g., Kashyap and Stein (2004) and Malherbe (2020)). This is why time-varying macroprudential capital requirements make sense.[2]

A key step in evaluating the policy tradeoff is understanding how capital regulation interacts with a bank's lending decisions. In this chapter, we first provide some institutional background and stylized facts about Basel III and the CCyBs. We then present a detailed discussion of the economics of bank capital requirements and lending, drawing on ideas from a companion working paper (Bahaj, Lattanzio, and Malherbe 2024). Our key research question is: how does a marginal change in capital requirements affect the aggregate supply of credit to the economy? Our approach is to use theoretical models to challenge typical priors, assess existing strategies seen in the empirical literature, and provide guidance for future research.

The are several takeaways from this chapter. First, most work on capital requirements and lending focus on what we refer to as a composition effect: from the bank point of view, equity capital is the more expensive form of finance, so shifting the composition of liabilities towards capital increases marginal cost and, hence, leads to a cut in lending. Second, taking this effect at face value, identifying its aggregate strength is challenging. Recent studies, have for instance made

[2] There seems to be a consensus among academics and policy makers that a constant capital requirement would unnecessarily amplify the business and financial cycles. Setting additional buffers during booms and loosening them in downturns is a way to correct such effect. This is where the countercyclical buffer terminology comes from (even though the buffers themselves are supposed to be positively correlated with the economic cycle, which makes them, strictly speaking, procyclical variables).

advances on identification issues such as the non-random assignment of regulation and how to deal with the spillovers associated with bank competition. However, the empirical literature has yet to tackle both simultaneously. Third, equity capital being more expensive on average does not mean that tighter capital requirements always raise the bank's effective marginal cost. One should therefore not take a negative lending response for granted. We highlight two effects that can lead to a positive lending response: first, a rationing effect when depositors value liquidity services (Begenau (2020)); and, second, a forced safety effect when deposits benefit from deposit insurance (Bahaj and Malherbe (2020)). Fourth, if bank assets are heterogeneous, different banks will value otherwise identical loans differently. Furthermore, changes in capital regulation can shift those relative valuations in ways that are complex and hard to measure. In our view, the thorny issues associated with heterogenous valuation of loans (or assets) by banks are among the most interesting and important open questions in this strand of literature.

2 Basel III and CCyBs

The 28 jurisdictions that form the Basel Committee, and comply to its regulation, cover most of the world's major economic and financial centres (including the US, EU and China).

Basel III prescribes a series capital requirements. For simplicity, we will mainly focus on capital requirements in terms of ratio of Tier 1 capital to risk-weighted assets.[3] Basel III specifies that all banks must adhere to a minimum 6% capital requirement at all times, plus an additional 2.5% capital conservation buffer that the bank can use temporarily to absorb a loss but will face dividend restrictions as a result. The minimum requirement and conservation buffer are homogeneous across jurisdictions and lending location, and are not time varying. Ignoring the subtleties of the capital conservation buffer, a stylized way to formalize this is to state that, at any date t, bank capital, which we denote e_t, must be at least a fraction $\underline{\gamma}$ of total risk weighted assets, which we denote x_t.

$$e_t \geq \underline{\gamma} x_t.$$

The CCyB

The constant requirement described above is supplemented by the countercyclical capital buffer (CCyB), which is time-varying by design. This buffer is additive to the static requirement. We can therefore formalize it as follows

[3] Tier 1 capital includes some contingent capital instruments beyond common equity that enable the bank to absorb losses on a going concern basis. The minimum capital requirement for common equity is 1.5 percentage points lower. The Basel III framework also contains an additional 2 percentage point requirement for Tier 2 capital, which can be thought of as gone-concern loss absorbing liabilities such as subordinated debt.

$$e_t \geq \left(\underline{\gamma} + CCyB_t\right) x_t, \qquad (1)$$

where $\text{CCyB}_t \geq 0$ is the additional buffer at date t.

However, since economic and financial cycles are not perfectly synchronous globally, CCyBs may vary across jurisdictions. Hence, equation (1) abstracts from banks transjurisdictional operations. Concretely, CCyBs are set at the discretion of regulators in a particular jurisdiction and apply to all bank loans made in the regulator's jurisdiction, irrespective of which jurisdiction the bank belongs to. Hence, it is a host-country rule.[4] Taking this into account, and considering a bank operating in two jurisdictions, we can write the constraint as:

$$e_t \geq \left(\underline{\gamma} + \text{CCyB}_t\right) x_t + \left(\underline{\gamma} + \text{CCyB}_t'\right) x_t'.$$

That is, a bank with risk-weighted assets x_t in a first jurisdiction and x_t' in the second one, faces a baseline capital requirement $\underline{\gamma}$ on both exposures, but faces potentially different additional buffers on exposures in specific jurisdictions (CCyB_t and CCyB_t', respectively).

A branch of the literature studies strategic interactions between national regulator who set capital requirements for banks that operate transnationally. Cross-border issues and strategic interactions between regulators is not the focus of this paper (for these, see, e.g., Dell'Ariccia and Marquez (2006) for the home-country rule, and Bahaj and Malherbe (2024) for the host-country one). So, in our analysis below, we simply consider constraints of the type

$$e_t \geq \gamma_t x_t,$$

where $\gamma_t \equiv \underline{\gamma} + \text{CCyB}_t$.

The use of CCyBs: Stylized facts

The countercyclical buffer regime of Basel III was phased-in between 1 January 2016 and year-end 2018 and became fully effective on 1 January 2019. Almost all countries had announced their initial rate (which could be zero) by the end of 2016.

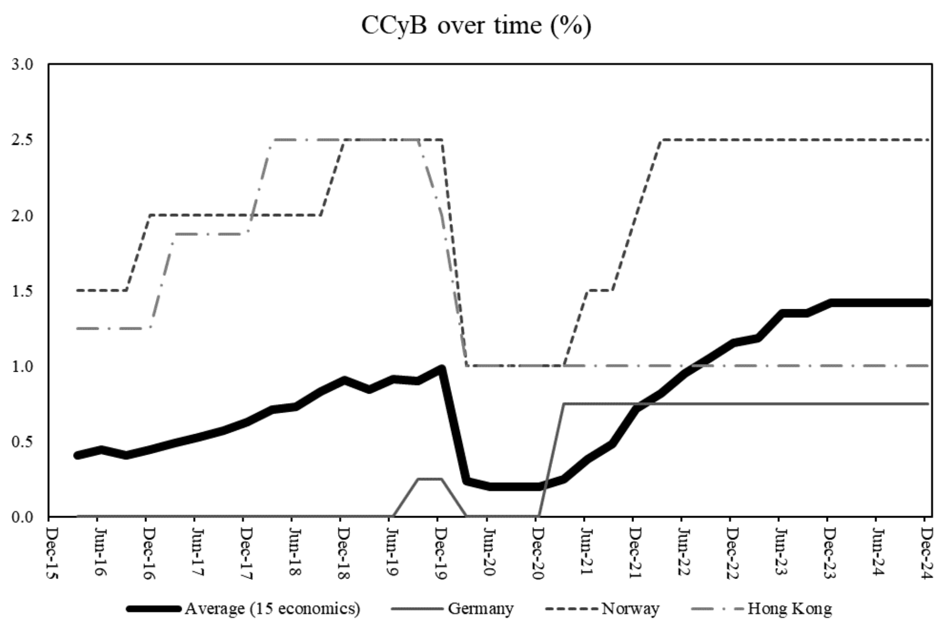
Some countries have never activated the buffer (that is, they have been setting $\text{CCyB}_t = 0$ at all t). However, by 2024, 15 of the 33 countries that report their CCyB level to the Basel Committee

[4] Even though, *de jure* the rule is home-country based, *de facto* it is a host-country rule. The regulator can always impose the requirement on banks within its jurisdiction (domestic banks and subsidiaries of foreign banks). For banks outside the regulator's jurisdiction that may lend within it, the framework is based on obligatory reciprocity: foreign regulators must impose the equivalent CCyB on the banks they regulate. For members of the Basel committee the obligation to reciprocate is encoded in the law governing capital regulation. Note, however, that reciprocity is only mandated up to a buffer of 2.5%; thereafter it is voluntary.

had set strictly positive buffers at some point, and many of them have been actively adjusting them over time. The passive countries (those who haven't activated the buffer yet) include very large economies, such as the US, China and Japan. The most active users of the CCyB tend to be smaller countries in Northern and Central Europe, but also include France, Germany and the UK as well as developed countries in East Asia, particularly Hong Kong and South Korea.[5]

For these countries, one can think of the CCyBs as being to the Financial Policy Committees what the short-term interest rate is the the Monetary Policy; they are use to set the stance of macroprudential policy. The history of using the CCyB is short but there have been two pronounced hiking cycles among members so far. The cross country average rate (among active users) rose from about 50 to 100 basis points between end 2016 and end 2019 before being cut to around 10 basis points at the outset of the Covid pandemic. Policymakers began hiking again in around mid 2021 and, at the time of writing, the average buffer was around 140 basis points. However, heterogeneity is substantial (if only given the number of countries that have yet to vary the buffer). The peak rate at a country level is 250 basis points a level hit both at the end of 2019 and by mid 2024. (See Figure 1 for an illustration).

Figure 1: *Sample of CCyB since 2016*



The thick black line only includes jurisdictions that have ever set a strictly positive CCyB.

---

[5] While Spain has not activated the CCyB, their system of dynamic (loan-loss) provisioning also aims at increasing capital buffers in good times. Evidence suggests that this has countercyclical effects (Jiménez et al. (2017)).

3 Lending responses: from theory to empirics to policy

Macroprudential regulators typically perceive themselves as trading off a safer banking system as a whole, with larger capital buffers to absorb losses, versus the contraction in aggregate lending that results from more equity financing. Estimating how steep such a contraction could be is therefore of first order importance. Given such policy decisions, by design, depend on the current and projected state of the economy and of the financial sector, this quest is fraught with endogeneity issues. Still, one can find in the literature a series of studies that exploit plausible sources of exogenous variation in order to overcome the identification challenges. The purpose of this section is to use theory (which we mainly draw from (Bahaj, Lattanzio, and Malherbe 2024)) to revisit popular strategies and assess whether, and to which extent, local estimates can be used as plausible proxies of aggregate responses.

We first outline a model of a single bank to illustrate our general approach. Then, we extend it to an equilibrium model of imperfect competition and finally turn to what the model implies for different identification strategies. A key assumption in this section will be that bank funding costs are exogenous: in particular, we will assume there is a constant excess cost of equity capital (over the cost of deposits). In the literature, this a popular way to capture various frictions (e.g., associated with moral hazard or stigma) that would affect the private cost of equity from a bank shareholder perspective (e.g., Hellmann, Murdock, and Stiglitz (2000)), or that deposits are implicitly subsidized by government guarantees (e.g., Thakor (1996)), or the tax advantage of debt. Such a reduced form approach has obvious benefits in terms of analytical tractability, but is far from innocuous in terms of results. We will elaborate on this in Sections 4 and 5. But, in the meantime, we will exploit the analytical convenience of this assumption to isolate some important mechanisms.

3.1 A baseline model of a bank's lending decision

To set the stage, consider a single bank that faces a downward sloping (inverse) demand function for loans. An amount of loans x generates an expected gross revenue $R(x)$, which is increasing, concave and bounded from above, and with $R_x(0)$ arbitrarily large. The bank starts with an infinitesimally small amount of inside equity. It can finance itself by raising deposits (denoted d), at a net interest rate normalized to zero, and with outside equity capital (denoted e), that requires an exogenous excess rate of return of $\rho > 0$. The bank's balance sheet reads

	Assets	Liabilities	
loans	x	e	capital
		d	deposits

The bank faces a capital requirement denoted γ. The associated constraint is

$$e \geq \gamma x.$$

Given $\rho > 0$, the constraint is binding at the bank optimum. Hence, $d = (1-\gamma)x$ and $e = \gamma x$. The bank's problem is to maximize the expected payoff to existing shareholders, that is

$$\max_x R(x) - \big((1-\gamma)x + \gamma x(1+\rho)\big).$$

The optimal amount of lending, x^*, is implicitly pinned down by the first order condition

$$\underbrace{R_x(x^*)}_{\text{marginal revenue}} - \underbrace{(1+\gamma\rho)}_{\text{marginal cost}} = 0. \quad (2)$$

This first order condition says, intuitively, that the bank equates the revenue from the marginal loan to the marginal cost of financing the loan, the latter depends on the excess cost of equity, ρ, and the proportion of equity finance, γ. Throughout, we will denote equilibrium objects with an asterix,*. Our main object of interest is the change in equilibrium lending (x^*) that is generated by a marginal change in capital requirement (γ). In short, we refer to it as *the lending response*. This response can be obtained by applying the implicit function theorem to the first order condition

$$\frac{dx^*}{d\gamma} = \frac{\rho}{R_{xx}(x^*)} < 0. \quad (3)$$

Here, the lending response depends on two factors only. The numerator, $\rho > 0$ captures the excess cost of equity capital. This factor reflects that an increase in γ forces the bank to finance itself with a costlier mix of funds as it forces a shift in the *composition* of liabilities away from deposits and towards equity capital. The bank's marginal cost goes up and, in turn, lending declines. What matters here is that, from the bank's perspective, the cost of equity is higher than that of deposits. If, instead of costly equity, there is an exogenous and constant unit subsidy to deposits, the result would be equivalent. What matters is the cost differential. We discuss economic mechanisms that can generate a difference between the cost of equity and deposits below. At this point, note the larger the differential, ρ, the stronger this *composition effect*.

The denominator in equation (3), $R_{xx}(x^*) < 0$ measures the steepness of the bank's marginal revenue function at the initially optimal level of lending. The flatter the marginal revenue curve the larger the adjustment in quantities that is needed to restore the equality between revenues and cost, and so the more sensitive the optimal quantity of lending is to changes in capital requirements.

The baseline case features a single bank and so follows a partial equilibrium approach. However, the econometrician (and the policy maker) is typically interested in aggregate consequences of changes in capital requirements. That involves modelling how banks interact with each other. In addition, empirical approaches often adopt difference-in-difference designs, for instance, in which some banks are affected by changes in capital requirements, and others not. Under perfect competition, treated banks would simply be priced out of the market. Hence, models that abstract from imperfect competition are not well suited to guide most empirical work on the topic.[6]

3.2 Equilibrium lending responses

We follow Schliephake and Kirstein (2013) and Lattanzio (2024) and adopt a Cournot approach of the loan market.[7] Imagine the economy has $N \geq 2$ banks like the one described above. They are ex-ante symmetric and face the same initial capital requirement γ. We assume revenue at the bank level can be written as $R(x) = xr(x + x')$ where x' is lending by other banks, so $X = x + x'$ and $r(X)$ is the aggregate marginal revenue function from lending.[8] In a symmetric Cournot equilibrium, the bank-level lending response to a change in γ can be written as

$$\frac{dx^*}{d\gamma} = \frac{\rho}{R^*_{xx}(x) - \left(\frac{N-1}{N}\right) r_x(X^*)}. \quad (4)$$

The difference with the baseline model is that the denominator explicitly accounts for the competitive environment and the fact that other banks are also affected by the capital requirement, which alters the marginal revenue curve at the bank level. Since equation (4) is evaluated at the Cournot equilibrium, its existence relies on the equilibrium elasticity of $r_x(X)$ being less than unity, that is $-\frac{X^* r_{xx}(X^*)}{r_x(X^*)} \equiv \eta^* < 1$. As this object turns out to be important in our analysis, with a slight abuse of terminology, we will refer to it as the elasticity of loan demand.

By definition, the aggregate lending response in a equilibrium is

---

[6] Assuming the cost of funds are constant (ρ for equity and 1 for deposits), the bank's marginal cost is also its average cost of fund. If the bank's average cost is higher than that of its competitors, it cannot survive without market power in the loan market. Alternatively, under perfect competition in the loan market, banks facing higher capital requirement than others could survive if they had market power in the deposit and/or equity market.

[7] Besides offering tractability, this approach has conceptual appeal. In practice, banks cannot adjust equity capital rapidly. Given banks face capital requirements, they effectively face capacity constraints on the quantity of loans they can make in the short term (Schliephake and Kirstein 2013). So, even if they compete a la Bertrand in the loan market, the Cournot equilibrium is the likely outcome (Kreps and Scheinkman (1983) made this general point of a Cournot equivalence with capacity constraints, and Schliephake and Kirstein (2013) and then Lattanzio (2024) applied it to the banking sector).

[8] The individual bank's problem is therefore $\max_x r(x + x')x - (1 + \gamma\rho)x$, where the bank takes x' as given. Taking the first order condition and imposing $x' = (N - 1)x$ in equilibrium, pins down x^* .

$$\frac{dX^*}{d\gamma} = N\frac{dx^*}{d\gamma}. \qquad (5)$$

This response is at the heart of the policy debate around capital regulation. As we described, macroprudential policy is typically calibrated to reflect a perceived trade off between a safer banking system as a whole, with larger capital buffers to absorb losses, versus the contraction in aggregate lending that results from more equity financing. This is why $\frac{dX^*}{d\gamma}$ is a key object of interest for the empirical literature. Unfortunately, estimating those is an exercise fraught with endogeneity problems. For instance, responses to changes in CCyBs cannot readily be used as these changes obviously depend on the current state of the banking-sector's balance sheet and on the regulators projections about future financial and real outcomes.

3.3 Empirical estimation

Many studies that empirically assess the effect of a change in capital requirements on bank lending consider a research design that features plausibly exogenous "between-bank variation" in capital regulation. The typical setting has one group of banks experiencing tighter regulation than another for reasons unrelated to their lending.[9]

Such a between-bank research design maps into our framework as follows: consider a situation where only a subset of banks is affected by an increase in capital requirements (recall they initially face the same γ). Specifically, let n be the number of banks affected by the change (the *treatment* group) and $N - n$ be the number of unaffected banks (the *control* group). It turns out that the aggregate lending response to a change in the capital requirement of treated banks only, which we denote γ^{Tr}, is

$$\frac{dX^*}{d\gamma^{\mathrm{Tr}}} = n\frac{dx^*}{d\gamma}, \qquad (6)$$

[9] An important identification concern is that regulatory changes are correlated shocks to bank-specific loan demand, i.e. changes in the function $r(x)$. The typical method for controlling for loan demand is to use the estimator proposed by Khwaja and Mian (2008) and make use of within borrower variation – i.e. looking at two different banks' lending to the same borrower. The assumption then is there is no assortative matching between banks and borrowers, and banks are not specialized lenders such that one bank is a good control for another. These are assumptions that the literature has questioned (see, for example, Chang, Gomez, and Hong (2023; Paravisini, Rappoport, and Schnabl 2023)). However, we leave the issue of how best to control for loan demand outside the scope of this paper and, for tractability, we will hold the demand function fixed.

where $\frac{dx^*}{d\gamma}$ is the lending response derived above (Equation (4)). So the aggregate lending change implied by an increase in the capital requirement for a subset of banks only is simply the number of treated banks times the bank-level change that would occur if all banks were treated.

This result does not mean, however, that each treated bank adjust lending by $\frac{dx^*}{d\gamma}$ and that control banks do not react. The reason is the following. Imagine that the treated bank response was $\frac{dx^*}{d\gamma}$. Then, control-group banks would effectively face less competition and would *expand*. But, then, treated banks would face more competition and contract further, and so on. What (Bahaj, Lattanzio, and Malherbe 2024) finds is that: i) The lending response of treated banks, $\frac{dx^{\mathrm{Tr}}}{d\gamma^{\mathrm{Tr}}}$, is larger (in absolute value) than $\frac{dx^*}{d\gamma}$; ii) Control-group banks expand, $\frac{dx^{\mathrm{Co}}}{d\gamma^{\mathrm{Tr}}} > 0$; and iii) The extra cut by the treated banks just offsets the expansion in the control group. So, the change in aggregate lending is simply proportional to the number of banks treated, which explains Equation (4). These effects have stark implications for identification.

Spillover bias

A research design based on a natural experiment that randomly assigns a group of banks, in a given market, to tighter policy would typically use a difference-in-differences estimator to account for confounding factors (examples abound, see for instance Imbierowicz, Kragh, and Rangvid (2018; Fraisse, Le, and Thesmar 2019; Arbatli-Saxegaard and Juelsrud 2022; Berrospide and Edge 2024)). Unfortunately, in a set up like ours, a difference-in-difference estimator (of $\frac{dx^*}{d\gamma}$) that compares treated and control banks in such an experiment suffers from spillover bias. The expression for how lending changes at a given treated bank versus a control bank, in response to a capital requirement change among treated banks, boils down to

$$\frac{dx^{\mathrm{Tr}}}{d\gamma^{\mathrm{Tr}}} - \frac{dx^{\mathrm{Co}}}{d\gamma^{\mathrm{Tr}}} \equiv \frac{dx^*}{d\gamma} + \underbrace{(N - \eta^*)\frac{dx^*}{d\gamma}}_{\text{spillover bias}}, \qquad (7)$$

where η^* is the equilibrium loan demand elasticity. This bias is positive (as $\eta^* < 1 < N$). The intuition follows directly from the mechanism described above: the contraction of lending by treated banks effectively reduces competition for control banks, which expand in response. Hence the difference between treated and control banks overstates the cut in lending that would occur if all banks were treated by tighter requirements. A regulator who extrapolates such estimates at the aggregate level would believe that tighter regulation generates an excessive cost in terms of the reduction in lending.

The spillover effect from the treated bank to the control bank constitutes a violation of the stable unit treatment value assumption (SUTVA) that is required for the use of a difference-in-differences estimator. While N is known in general, η^* is not directly observable, so one cannot directly rescale the estimate to eliminate the bias. The finance literature has proposed some solutions to deal with this problem. One option is to assume that demand curves are linear (Berg, Reisinger, and Streitz 2021). Then $r_{xx}(X^*) = 0$ and so $\eta^* = 0$. The spillover term can be accounted for by simply dividing the estimator by $1 + N$. However, Huber (2023) makes the point that non-linearities abound in banking. So, assuming linear demand to correct for spillovers is not a silver bullet. Specifying a functional form for η^* is another solution but estimating the elasticity of loan demand presents an additional set of identification challenges.

The literature has instead explored settings with multiple cross sections with heterogeneity in the fraction of treated units (e.g., Berg, Reisinger, and Streitz (2021) or Huber (2023), but also Mian, Sarto, and Sufi (2023)). Imagine a set of countries that are identical in size and have the same number of banks. These countries differ only in the fraction of banks that have been treated with tighter capital requirements. Assuming banks do not compete across countries, the heterogeneity in the number of treated banks can be used to identify η^*. Such an idealized setting, unfortunately, is unlikely to exist. Even if it was possible to obtain random variation in the number of treated banks, the industrial organisation of the banking sector vary dramatically across countries, leading to heterogeneity in η^*.

Random assignment

Setting aside spillovers, the typical concern with a difference-in-differences approach when assessing the impact of capital regulation is that random assignment may not be fully plausible. Take the example of a regulatory change that only affects banks that are above a given size threshold. A possible identification strategy is to assume that the exact value of the threshold is random. One can then compare banks that are just either side of it (see, e.g., Bassett and Berrospide 2018; Favara, Ivanov, and Rezende 2021). Critics of such an approach could argue that other shocks that affect bank lending could vary systematically with bank size, and that regulators are likely to be very careful on where to set the threshold. If this is the case, treatment cannot be interpreted as random, even just around the threshold.

To address the issue of non-random assignment along the size dimension, Gropp et al. (2019) compare similar banks across different markets with different exposures to capital regulation. In 2014, a European Union regulatory change raised capital requirements on the largest banks of each domestic market. The threshold was such that 50% of the market for deposits in each member state was covered. Given markets themselves have different sizes, the threshold in terms of bank size was effectively different across countries. This means that, across countries, there were banks of similar size that were either treated (because the size threshold is lower in their country) or not.

Such a design addresses the issue of random assignment but raises thorny issues with the interpretation of the coefficient. The difference-in-differences estimator that corresponds to the across-market control group strategy is also biased. The issue here is not spillovers between treated and control group, but that they each face spillovers that are inherent to the specific market they operate in. They should therefore not be expected to respond in a comparable manner to shifts in regulation that have sector wide effects, or to common confounding factors.

4 Endogenous marginal cost of capital (one of two)

We now turn to cases in which the bank's marginal cost of funds is an equilibrium outcome.

First, we show that introducing an upward sloping aggregate supply for equity capital leads to results that extend those of our baseline model in an intuitive way: Bank lending still decreases with capital requirements but the strength of the lending response also depends on the equilibrium elasticity of equity capital supply.

Second, we use our framework to revisit models that feature an endogenous cost of deposits. As we shall see, these can deliver results that stand in stark contrast with our baseline: the lending response can be *positive*.

4.1 Upward sloping supply of equity capital

As in our baseline model, the cost of deposit is 1. However, there is now an aggregate supply curve of equity, so that the value of ρ is determined in equilibrium and is increasing and weakly convex in K, the total equity capital raised by the banking system. Here, Banks act competitively in the equity and loan markets and therefore take $\rho(K)$ and $R_X(X)$ as given.

Accordingly, the representative bank now solves

$$\max_{x} xR_X(X) - (1-\gamma)x - \gamma x\big(1+\rho(K)\big),$$

Assuming that $R(.)$ satisfies conditions for an interior maximum for the bank's optimal amount of lending $x^*(\gamma)$. The equilibrium is pinned down by the following condition

$$R_X(X^*) - \big(1+\gamma\rho(K^*)\big) = 0,$$

where $K^* = \gamma X^*$. Hence the aggregate lending response is

$$\frac{dX^*}{d\gamma} = \frac{\rho(K^*) + K^*\rho_K(K^*)}{R_{XX}(X^*) - \gamma^2\rho_K(K^*)} < 0. \qquad (8)$$

Even though the equation for the lending response is more involved than that in the baseline (3), the fundamental logic does not change. There is an excess cost of capital. So, increasing the capital requirement increases the marginal cost of funds, and a lending cut ensues. The intensity of the lending response depends on the value of the excess cost, but also on its slope and curvature, as well as on the curvature of the loan demand function, as was the case before. Therefore, the main insight from this specification is that, if the banking system faces an increasing (and weakly convex) cost of raising equity, then tightening regulation when banks already need to raise equity (for instance, in response to large losses), will amplify the resulting cut in lending. This underpins much of the logic behind countercyclical regulation, as in Kashyap and Stein (2004) for example.

4.2 Upward sloping supply of deposits

In our baseline model and the various extensions, we have studied so far, the capital requirement binds because equity capital is intrinsically costly ($\rho > 0$). But another reason for it to be binding could be deposits are intrinsically cheap. This could be the case, for instance, if deposits provide liquidity services, if banks have market power over depositors, or if deposits benefit from underpriced government guarantees. Recent literature has looked at how such mechanisms interact with capital requirements. In this subsection, we mainly focus on the liquidity-service case. That of government guarantees is the focus of the next section.[10]

Assume that households obtain a non-pecuniary liquidity service from their deposit holdings that is a function of the aggregate stock of deposits, D. Let $\delta(D) \geq 0$ denote the liquidity service households obtain from the marginal deposit. Assume further that households value deposits more, at the margin, when they are scarcer, $\delta_D(D) < 0$. Normalising the opportunity cost of funds in the economy to 1 and assuming $\rho = 0$, a representative bank solves:

$$\max_x xR_X(X) - x\Big((1-\gamma)\big(1-\delta(D)\big) + \gamma\Big),$$

where $1 - \delta(D)$ is the unit cost of deposits, and $\delta(D)$ is taken as given by the bank. Assuming the second order condition holds, equilibrium aggregate lending, X^*, is implicitly pinned down by the zero-profit condition

$$R_X(X^*) - \big(1 - (1-\gamma)\delta(D^*)\big) = 0.$$

Given the capital requirement binds $(1-\gamma)X^* = D^*$, and we get:

$$\frac{dX^*}{d\gamma} = \underbrace{\frac{\delta(D^*)}{R_{XX}(X^*) + (1-\gamma)^2\delta_D(D^*)}}_{<0 \text{ (composition effect)}} + \underbrace{\frac{D^*\delta_D(D^*)}{R_{XX}(X^*) + (1-\gamma)^2\delta_D(D^*)}}_{\text{rationing effect}} \gtreqless 0.$$

As it turns out, the aggregate lending response combines two effects and is ambiguous in sign. The first term is strictly negative: for a given unit cost of deposit, an increase in γ increases the bank's cost of funds (this is the, now familiar, composition effect). The second effects captures the key mechanism of Begenau (2020): given an aggregate amount of lending, a tighter capital requirement

[10] There also is a new generation of papers studying the deposit franchise (the net present value of the rents that banks can extract in the deposit market) and its implication for bank risk management and financial stability (see, e.g., Drechsler, Savov, and Schnabl (2021) and DeMarzo, Krishnamurthy, and Nagel (2024), for interest rate risk management, and Döttling (2023) for the interactions with capital requirements around the zero lower bound).

means the banking sector can issue less deposits, effectively rationing liquidity supply.[11] Scarcer deposits are cheaper to issue (the more satiated the market for deposits, the less liquidity service they provide at the margin). This rationing effect makes lending more appealing at the margin and will dominate if the liquidity demand elasticity is greater than unity:

$$\frac{dX^*}{d\gamma} > 0 \Leftrightarrow -\frac{D^*\delta_D(D^*)}{\delta(D^*)} > 1.$$

Intuitively, if the deposit rate is not very sensitive to a change in the aggregate deposits, the composition effect prevails. Still, the rationing effect is one of the mechanisms that can lead to a positive lending response, as is the case in the economy calibrated by Begenau (2020).

Empirical considerations

The empirical literature so far, to our knowledge, has not attempted to disentangle this rationing effect from the composition effect and has therefore not identified cases where the lending response was be positive. This may be because such a general equilibrium mechanism (as emphasized in Begenau (2020)) is challenging to identify through research designs based on bank-level variation as described above. However, if one respecifies the problem as one where banks have market power, and have their own supply curve for deposits, a rationing effect also emerges. Unfortunately, this effect would also be associated with spillovers, linking back to the issues we discussed in the previous section.

Deposit rationing and central bank reserves

An issue that could limit the empirical relevance of the rationing effect is that the absence of bank reserves in the model above is not innocuous. This is because their absence hard-wires the link between equilibrium quantities of loans and deposits.

Imagine, instead, that the bank can also hold cash, or central bank reserves, denoted m, that pay a unit return and carry a zero risk-weight (as per the Basel regulation). In such an extension, the capital requirement is still biding at $e = \gamma x$, but the balance sheet identity becomes

$$e + d = x + m,$$

where d denotes deposits at the bank level. In this case,

$$d = (1 - \gamma)x + m,$$

so the loan quantity no longer pins down the amount of deposits. As a result, if competitive banks can issue deposits at a discount, and park them at the central bank where they earn a higher rate (or invest them in treasuries, for instance), they will do so. Hence, in a competitive equilibrium, banks will compete away the liquidity premium: demand for liquidity in the deposit market will

[11] See (Heuvel 2008) for an analysis of associated welfare costs.

be satiated. As a result, the rationing effect will not operate (and the composition effect will operate as long as equity capital carries an excess cost as in Section 3).

In practice, banks do have market power in the deposit market and there arguably are limits to the aggregate supply of risk-free claims that carry a zero risk weight.[12] However, despite this, it is difficult to argue that it is the (risk-weighted) capital requirement that constrains banks ability to provide liquidity services. For example, the period around the Covid Pandemic saw large fluctuations in D in many countries (mainly driven by increases in the aggregate supply of reserves) without corresponding increases in bank capital. Having said that, to our knowledge, there is no explicit empirical evidence, one way or another, on whether a link exists between the supply of liquidity services to the real economy and (risk-weighted) capital regulation.

The caveat 'risk-weighted' in the sentence above is an important one. If capital requirement is not risk-weighted (i.e. $e = \gamma(x + m)$), as is the case for the Basel III Supplementary Leverage Ratio in some jurisdictions, then the rationing effect is still present. Tighter leverage ratio requirements will constraint bank's ability to close the arbitrage opportunity between the return on the safe assets and the cost of deposits. While the present chapter focuses on risk-based capital regulation, it is worth mentioning that several papers present evidence that the leverage ratio limits the ability of banks to close risk-free arbitrage opportunities in financial markets (see, for example, Du, Tepper, and Verdelhan 2018; Hanson, Malkhozov, and Venter 2024).

5 Endogenous marginal cost of capital (two of two)

So far, we have ignored distortions associated with the risk of bank default. We now extend the analysis in such a direction.

5.1 The Merton decomposition

Let us go back again to the simple bank described in Section 3. The bank faces a binding capital requirement, and deposits pay a zero interest rate. But now, the reason why the deposit rate is fixed is that deposits are insured by the government (for simplicity, at no cost). For the moment, fix lending to $x = 1$, and assume that *realized* revenue, R, is stochastic. (Note that in the previous sections, R denoted *expected* revenue.) The bank also starts with an infinitesimal amount of inside equity and raises outside equity from investors who must break even in expectation. They have an opportunity cost of funds of 1. Denote $v(R)$ the dividend paid to investors at the end of the unique period.

[12] Under the Basel regulation, not only claims on central banks, but also claim on central governments (and associated entities) of OECD countries carry a zero weight.

	Assets	Liabilities	
loans	1	$e = \gamma$	capital
		$d = 1 - \gamma$	deposits

The bank will default on deposits if $R < 1 - \gamma$. Given shareholders are protected by limited liability, expected payoff to inside equity is

$$\pi = \mathbb{E}[R - (1-\gamma) - \gamma v(R)]^{+}, \qquad (9)$$

where $\mathbb{E}[.]^{+} \equiv \mathbb{E}[\max\{0, .\}]$. The investors' break even constraint reads

$$\mathbb{E}[\gamma v(R)] = \gamma, \quad (10)$$

substituting this breakeven condition into equation (9), and rearranging, allows for an intuitive decomposition of the payoff to inside equity holders (Merton 1977):

$$\pi = \underbrace{\mathbb{E}[R - (1-\gamma)] - \gamma}_{=E[R]-1 \text{ (NPV)}} + \underbrace{\mathbb{E}[(1-\gamma) - R]^{+}}_{\text{Implicit subsidy}}, \qquad (11)$$

The first term is expected revenue from the (unit) loan relative to the cost of funds in the economy (which is 1): i.e., it is the net present value of the bank's investment. The second term is a subsidy arising from deposit insurance. The decomposition makes clear that this subsidy accrues to the holders of inside equity. Both outside equity investors and depositors break even in expectation. But when the bank defaults (i.e., when $R < 1 - \gamma$), depositors break even at the taxpayer's expense. This implicit subsidy from the taxpayer is the expectation, over the default states, of the shortfall in asset value compared to the promised repayment to depositors (by definition, in the default states $1 - \gamma > R$, so the shortfall is positive). If insiders had unlimited liability, this expectation is the amount they would have to pay to make depositors whole. Given limited liability, they do not have to. And with deposit insurance, depositors do not demand compensation for the risk they will not be repaid in full as the taxpayer makes them whole.

For simplicity, we formalized the point above with equity capital. However, there is nothing specific here to equity. Imagine inside equity holder would issue junior bonds (or subordinated debt, or so-called AT1s). As long as such instruments do not benefit from government guarantees, that is, as long as they do constitute loss-absorbing capacity, substituting the associated break-even condition will lead to Equation (11) and the Merton decomposition will apply.

5.2 Excessive risk taking and capital requirements?

Government guarantees distort bank decisions towards excessive risk taking. Capital requirements can help mitigate the issue. We discuss two simple cases; one in which it is the composition of credit that is distorted and one in which it is the quantity.

Composition of credit: Asset substitution

Now assume $f(R)$ follows a uniform distribution with support $[1-\sigma, 1+\sigma]$, with $\sigma > \gamma$. The implicit subsidy as a function of γ and σ reads:

$$S(\gamma,\sigma) \equiv \underbrace{\left(\frac{\sigma-\gamma}{2\sigma}\right)}_{\text{default probability}} \times \underbrace{\left(\frac{\sigma-\gamma}{2}\right)}_{\text{expected shortfall given default}} = \frac{(\sigma-\gamma)^2}{4\sigma}.$$

The first thing to note is that the subsidy is increasing in σ. Ceteris paribus, bank profit increases with the riskiness of its asset. This is because shareholders get the upside while the taxpayer foots the bill in the downside. So, if the bank can choose between two loans with different σ, it will pick the one with the largest ((Kareken and Wallace 1978)). This is a simple case of the risk-shifting problem that is pervasive in the banking literature.[13] Linking back to Merton (1977), the implicit subsidy can be interpreted as a put option. Given limited liability, when the value of the bank is negative (i.e. in the default states), the shareholders can sell the bank at a strike price of zero. Naturally, the value of such option is increasing in the volatility of the underlying asset (more volatility means the option will more often be in the money).

Second, the subsidy is decreasing in γ. A higher capital requirement reduces taxpayer exposure (and a larger loss-absorbing capacity means that the put option is further out of the money). And, third, the cross-partial derivative of the subsidy is negative too: the benefit to the bank from an increase in σ decreases with γ. This can be interpreted as larger capital buffers mitigating risk-shifting incentives.[14]

Quantity of credit: Risk-shifting and over lending

To see how deposit insurance can distort the quantity of credit, consider our baseline model, with choice variable x, but with risk, default, and limited liability.

---

[13] Several studies provide evidence that deposit insurance incentivize banks to take on more risk (e.g. Keeley (1990), Anginer et al. (2014) Calomiris and Jaremski (2019), Calomiris and Chen (2022)), especially when controlled by shareholders (Leaven and Levine, 2009). Cucic et al. (2024) show that an increase in the deposit insurance coverage led to an increase in credit to riskier firms. However, the effect is entirely driven by the reallocation of deposits across banks, rather than by changes in risk-taking behaviour within banks. Danisewicz et al. (2022) found no evidence insured banks are less sound.

[14] However, to fully eliminate such incentives, we need: $\gamma > \sigma$ so that the bank is completely safe.

Assets		Liabilities	
loans	x	$e = \gamma$	capital
		$d = 1-\gamma$	deposits

Here we assume the date-1 realized revenue from lending is $AR(x)$, where $A \sim f(A)$ is a positive stochastic variable with $\mathbb{E}[A] = 1$, and assume the needed regularity conditions hold for all problems being well behaved. The bank defaults if

$$AR(x) < (1-\gamma)x.$$

So, we can define a default boundary realisation of A as $A_0(x) \equiv \frac{(1-\gamma)x}{R(x)}$. (For readability, we henceforth omit dependencies on x for endogenous objects and their derivatives, and use again an asterix to indicate when a function is evaluated at the equilibrium level of lending). Using the Merton decomposition the problem of the bank is

$$\max_x \mathbb{E}[R] - x + \int_0^{A_0} \big((1-\gamma)x - AR\big) f(A) dA$$

with associated first-order condition

$$\mathbb{E}[R_x^*] - 1 + \underbrace{\int_0^{A_0^*} \big((1-\gamma) - AR_x^*\big) f(A) dA}_{\equiv S_x(x^*,\gamma)} = 0.$$

One can interpret the term S_x^* as how much the marginal loan affects the implicit subsidy from deposit insurance. Note that $S_x^* \geq 0$, strictly if the bank defaults with strictly positive probability. This means that the marginal loan is effectively subsidized and the bank will finance negative NPV loans as the downside risk is born by the taxpayer; another instance of risk-shifting.

Now, the lending response is

$$\frac{dx^*}{d\gamma} = \frac{S_{x\gamma}^*}{-(R_{xx}^* + S_{xx}^*)}, \qquad (12)$$

and its sign is that of $S_{x\gamma}^*$. In this case, an increase in γ unambiguously reduces the marginal subsidy: $S_{x\gamma}^* < 0$. Hence, like in Section 3, the bank responds to an increase in capital requirement with a cut in lending. The difference is that in Section 3, the numerator was an exogenous constant (ρ) and here we have an endogenous object $S_{x\gamma}^*$. This object is generally complex and is studied at great length in Bahaj and Malherbe (2020). A key point to recognize is that in the present example it is unambigious that $S_x^* > 0$ and $S_{x\gamma}^* < 0$. Hence, banks have an incentive to overlend, but this incentive can be mitigated through tighter capital regulation. However, as we shall see, these results do not necessarily hold in a more general setup: i) The marginal subsidy can be negative (which leads to *underlending*); ii) And, perhaps more importantly, the marginal subsidy can be *increasing* in the capital requirement, which leads to a positive lending response.

5.3 Guarantee overhang and positive lending responses

A simple example of a case with a positive lending response is to consider a bank with a balance sheet that includes assets with imperfectly correlated payoffs. These could be legacy loans or other inframarginal loans. Imagine the marginal loan yields a gross revenue R_1 following a distribution with mean $\mu_1 > 1$. Given the other assets on the balance sheet, the bank either survives or defaults.[15] The bank will finance the additional loan if and only if

$$\underbrace{p\mathbb{E}[R_1 - (1-\gamma) \mid \text{Survival}]}_{\text{exp. residual cash flow accruing to shareholders}} \geq \underbrace{\gamma}_{\text{extra capital needed}}$$

The term inside the expectation on the left-hand size is the realized residual cash flow on the marginal loan. This is how much is left from the revenue associated with the marginal loan, after deducting the associated repayment to depositors. In the states where the bank survives, which occurs with some probability p, this residual cash flow accrues to the shareholders. For the bank to decide to make the marginal loan, intuitively, the expected residual cash flow accruing to shareholders needs to exceed the capital used to finance the loan.

Using the Merton decomposition, this condition can be rewritten as

$$\underbrace{\mu_1 - 1}_{\text{NPV}} + \underbrace{(1-p)\mathbb{E}[(1-\gamma) - R_1 \mid \text{Default}]}_{\equiv S_1(\gamma)} \geq 0$$

where S_1 can also be interpreted as a marginal subsidy: it is the increment to the bank's implicit subsidy if the bank finances the marginal loan.[16] The marginal subsidy captures that the residual cash flow accrues to the taxpayer in the states where the bank defaults.

Now, imagine that $\mu_1 > 1$ and the return on the marginal loan is independent of that of the rest of the bank's assets (in particular imagine $\mathbb{E}[R_1 \mid \text{Default}] = \mu_1$). The financing condition becomes

---

[15] That the marginal loan is infinitesimally small allows us to ignore that financing an extra loan may affect the probability of bank default. This simplification simply allows us to zoom in the effect of γ on bank default.

[16] The term in S_1 is also similar to a funding value adjustment as discussed by Duffie, Andersen, and Song (2019). The main distinction between the marginal subsidy and a funding value adjustment is that the latter arises from preexisting debt, not a guarantee, and reflects a transfer to/from exiting debtholders rather than the taxpayer. Economically, the two have very similar effects, however.

$$\underbrace{\mu_1 - 1}_{\text{NPV}} - \underbrace{(1-p)\big((\mu_1 - 1) + \gamma\big)}_{>0} \geq 0. \quad (13)$$

So, the bank values the loan lower than its net present value. When the bank defaults, not only does it not capture the positive NPV, but it also makes an additional equity capital loss. If the probability of bank default is non negligible the financing condition will be violated: The positive NPV loan will not be taken up, even though it could be mainly financed by insured deposits. This is an instance of what Bahaj and Malherbe (2020) dub the *Guarantee Overhang*: if the existing balance sheet is risky, banks may pass on positive NPV loans because they do not fully internalize the associated surplus.

Consider a case where, given γ, the bank prefers not to make the loan. Now imagine regulation tightens. As γ increases, p increases. As p tends to 1, the second term in (13) vanishes. The bank fully internalizes the surplus and decides to make the loan (since $\mu_1 > 1$). An increase in capital requirement generates a *positive lending response*.

In our example, the loan carries purely idiosyncratic risk. The result doesn't hinge at all on this. The key to a positive lending response is heterogeneity in residual cash flows. This means that when the bank defaults, the marginal loan can still generate positive residual cash flows for shareholders, even if, on average, the bank's assets are underwater. This heterogeneity can come from many different sources. Imperfect correlation between the loan return and that of the rest of the bank portfolio is one. But heterogeneous residual cash flows can also arise among loans that are perfectly correlated (for instance, if they face different risk weights or if their returns have identical distribution functions, but with different means).[17] In reality, residual cash flows will always exhibit some degree of heterogeneity. In our risk-shifting example in Subsection 5.2, the residual cash flows on the marginal loan are always less than those on the average loan, so the marginal subsidy is always (weakly) positive and decreasing in the capital requirement. As (Bahaj and Malherbe 2020) show, many standard assumptions in the literature, made for the sake of tractability, also rule out relevant cash flow heterogeneity and therefore bias results towards always finding negative lending responses.

[17] One specific case of where heterogeneity in residual cash flows can arise is geographical variation in lending. For example, Puri, Rocholl, and Steffen (2011) show that state banks in Germany that were heavily invested in US subprime loans cut back on loans to German retail borrowers during the 2007 to 2009 financial crisis. Given the relative safety of German retail borrowers a reasonable interpretation is that these loans were positive NPV, but the residual cash flows on them had a very different distribution from the cash flows from US subprime lending. Ex-post, the increased losses from the US worsened the guarantee over hang problem and made German banks less reluctant to lend at home. A higher capital requirement may have prevented this effect.

In a framework with legacy loans and x as the continuous choice for the amount of new loans (the main model in (Bahaj and Malherbe 2020)), the formula for the lending response is again given by equation (12). However, $S^*_{x\gamma}$ can be decomposed as

$$S^*_{x\gamma} = \underbrace{-(1-p^*)}_{\text{composition effect}} + \underbrace{p^*_\gamma \mathbb{E}\left[AR^*_x - (1-\gamma) \mid \overbrace{A = A^*_0}^{\text{Default Boundary}}\right]}_{\text{FSE}}. \quad (14)$$

The first term is the standard composition effect, which we already had in the model with exogenous cost of capital (strictly speaking, the composition effect should also be scaled by denominator in equation (12)). It captures that higher γ puts more capital at risk, which tends to make the lending response negative.

The second term on the RHS, is called the *Forced Safety Effect* (FSE). By forcing the bank towards safety tighter regulation alters the states of the world where the bank survives and hence how it values the marginal loan. The term p^*_γ is the change in the probability of default associated with the change in γ. So, it captures the shift in the probability of default. This shift makes the bank internalize residual cash flows that would otherwise accrue to the taxpayer. In particular, $\mathbb{E}[AR^*_x - (1-\gamma) \mid A = A^*_0]$ is the expected residual cash flow on the marginal loan, evaluated at the states where the bank is on the boundary of default. If this object is negative (as in the example of Subsection 5.2), the FSE is negative and this reinforces the composition effect. However, if the object is positive (as in the example with idiosyncratic risk), the FSE is positive and at least partially offsets the composition effect. In some cases, it dominates, and the lending response is positive.

In reality, the returns to new lending opportunities are likely to be correlated to that of the bank existing portfolio. However, they will inevitably have an idiosyncratic component. And the key message from this section is that the sign of the lending response depends in a complex way on the composition of credit.[18]

5.4 Bank specialisation and the composition of credit

An important takeaway from the analysis above is that, given an existing balance sheet, banks will overvalue investment opportunities that generate negative residual cashflows in states where its

[18] As noted in Bahaj and Malherbe (2020): *'Banks that have a high risk of failure are likely to have a negative lending response because of the strength of the composition effect. One implication of this prediction is that if bankers argue that higher capital requirements would lead to a substantial decrease in lending, they must believe that the composition effect, and hence default probabilities, are large. This would mean that banks are receiving substantial subsidies in the first place.'*

overall portfolio is in the red.[19] And, conversely, they will undervalue investment opportunities that generate positive residual cash flows in these states. This naturally points towards an incentive for banks to specialize in downside risk.[20] Harris, Opp, and Opp (2023) offers an elegant formalisation of this phenomenon. In a set up where a continuum of perfectly competitive banks face capital requirements and lend to an heterogenous continuums of firms, they show how, endogenously, banks perfectly specialize in equilibrium, and how an aggregate demand for bank equity capital emerges. In such a set up, changes in the capital requirement can have dramatic, and sometimes counterintuitive effects on the equilibrium composition of credit.[21]

In practice, it is impossible for banks to perfectly specialize, especially if they want to be large. Banks typically make loans to multiple types of borrowers. Nonetheless, changes in the capital requirement alter the composition of credit. Consider a bank that specializes in two types of loans, with separate revenue functions, each similar to what we have used above. We denote x_1 and x_2 the quantity of lending of each type, so that the total revenue function is: $A\big(R(x_1) + R(x_2)\big)$. In addition, let us introduce risk-weights. The bank still face a capital requirement, but loan amounts are now weighted by α_1 and α_2, respectively. So, the capital requirement constraint reads

$$e \geq \gamma(\alpha_1 x_1 + \alpha_2 x_2).$$

Defining $\gamma_1 \equiv \gamma\alpha_1$ and $\gamma_2 \equiv \gamma\alpha_2$ makes clear that applying different risk-weights is isomorphic to applying different capital requirements. That is, the constrain can equally be written

$$e \geq \gamma_1 x_1 + \gamma_2 x_2,$$

which also make apparent that our present set up can be interpreted in terms of differential sectoral capital requirements.

Now, let us assume initial symmetry ($\gamma_1 = \gamma_2$) and that the initial optimum for the bank is an interior solution $x_1^* = x_2^*$. Consider a small change in γ_1. For loans of type 1, we get:

$$S^*_{x_1\gamma_1} = -(1-p^*) + p^*_{\gamma_1}\mathbb{E}\big[AR^*_{x_1} - (1-\gamma) \mid \text{Default Boundary}\big].$$

---

[19] Put differently, a bank will overvalue new loans that are similar to legacy loans. Landier, Sraer, and Thesmar (2015) provide evidence of this behaviour from a precrisis U.S. subprime lender.

[20] Malherbe and McMahon (2024) show how the use of derivatives to double down on downside risk can exacrebate the overlending problem introduced above. In this context, one can think of derivatives as tools used to reduce residual cash flow heterogeneity in order to maximize the value of the put option associated with limited liability.

[21] Along the same lines, Oehmke and Opp (2023) show that using differentiated capital requirements to support a green transition (i.e., setting capital requirements higher for brown firms than green firms) can generate ambiguous responses, including a decrease in green lending.

The situation is very similar to the generalized result above: loans of type 1 have to be financed with more capital, which generates the composition effect. In addition, there is a Force Safety Effect, an increase in γ_1 makes the bank safer so that it internalizes the residual cash flows along the default boundary ($p^*_{\gamma_1}$ is the increase in survival probability induced by the change).

However, for loans of type 2, we have:

$$S^*_{x_2\gamma_1} = p^*_{\gamma_1}\mathbb{E}\big[AR^*_{x_2} - (1-\gamma) \mid \text{Default Boundary}\big].$$

There is no composition effect here since γ_2 has not changed. Still, the change in γ_1 triggers a forced safety effect. Given initial symmetry, we have $R_{x_1}(x^*_1) = R_{x_2}(x^*_2)$ and $R_{xx}(x^*_1) = R_{xx}(x^*_2)$. Consider the following difference-in-differences comparison

$$\frac{dx^*_1}{d\gamma_1} - \frac{dx^*_2}{d\gamma_1} = \frac{S^*_{x_1\gamma_1} - S^*_{x_2\gamma_1}}{-(R^*_{xx} + S^*_{xx})} \equiv \frac{-(1-p^*)}{-(R^*_{xx} + S^*_{xx})}. \quad (15)$$

This calculation isolates the composition effect. However, it filters out the forced safety effect.

This result as implications for empirical work using within bank variation in capital requirements. Behn, Haselmann, and Wachtel (2016) use a randomly assigned change in risk weights in Germany during the 2008-2009 financial crisis to study, in a difference-in-differences research design, how banks react when a portion of their loan portfolio is treated by tighter requirements relative to loans that are not. Given random assignment, the treated and control groups should be comparable which is equivalent to saying $x^*_2 = x^*_1$. Using the estimator above as a proxy for the lending response induced by the change misses the forced safety effect.

This missed effect is another violation of SUTVA. However, this time the spillover is not between banks but within them. Loans of type 2 are not directly treated by tighter regulation but there is a spillover through the overall riskiness of the bank's balance sheet. Any regulatory that affects bank default probability, will make the bank re-price all marginal loans, regardless of whether the loan is directly affected or not. If a regulator substantially restricts risk-taking in one line of business, this will cause the internalisation of the residual cash flows of other of business lines in more states of the world. If those cash flows are positive (perhaps because the alternative line of business generates relatively safe returns), the bank will expand in that dimension. This prediction is consistent with evidence in Acharya et al. (2018), who finds that when the Central Bank of Ireland imposed restrictions on the issuance of risky loans to urban borrowers, banks that were initially heavily exposed aggressively expanded their issuance of loans to safer borrowers in rural counties.

Overall, these spillovers among loans complicate empirical work comparing loans within a bank's balance sheet. Even more daunting, however, is that the between-bank analysis described in Section 3 would be affected in complex and ambiguous manner if banks have heterogeneous balance sheets. For instance, a comparison of two different banks' lending to the same borrower (for example, in a standard Khwaja and Mian (2008) setting) is hard to interpret if heterogeneity in the rest of the balance sheet causes the two banks to value the loan differently.[22]

6 Conclusion

Conventional wisdom has it that banks cut lending following an increase in capital requirements. In this chapter, we have delved into the economics of the relationship between capital requirements and lending. Our model starts from first principles and is based on a logic that can be applied to many situations: derive the first order condition to obtain lending as an implicit function of the capital requirements, and from there the lending response can be derived from the implicit function theorem. We showed that our baseline model and many variations of it deliver a prediction in line with the conventional wisdom: the lending response is negative. This reflects a composition effect. If equity capital is a more expensive source of funds for a bank at the margin, then the switch in the composition of funds raises marginal cost and results in a decline in lending. Analysing this effect can help think formally at the drivers of the magnitude of such cuts at the aggregate level. In doing so, we have highlighted how imperfect competition, the presence of confounding factors, and heterogeneity among different types of assets raise substantial hurdles for empirical estimation that the literature has yet to fully overcome.

Moreover, we have shown that the composition effect is not the end of the story. The recent literature has emphasized that banks have large deposit franchises and depositors are often willing to accept low returns for their holdings in exchange for liquidity services. If tighter regulation constrains deposit supply this can generate a rationing effect that lowers the cost of deposits and has the potential overturn a negative lending response. A positive lending response can also emerge when accounting for bank default and the distortion generated by government guarantees. Then lending responses depend in complex ways on the joint distribution of residual cash flows of the existing bank portfolio and of the residual cash flows of lending opportunities. We highlighted the force safety effect that can either reinforce or offset an otherwise negative lending response. Overall, we hope we have convinced the reader that the economics of bank capital requirements and lending are much more subtle and complex than what conventional wisdom suggests.

In fact, even though we have studied a broad range of mechanism, a crucial dimension is missing: dynamics. In reality, banks do not adjust instantaneously to changes in regulation. Empirical evidence suggests that they take a few quarters to do so, and that margin of adjustment depend on the state of the economy (Bahaj et al. 2016). But spillover effects are even more complex in a

[22] See Internet Appendix 2 in Bahaj and Malherbe (2020) for an example for how changes in the initial heterogeneity in banks' balance sheets can have non-linear and unpredictable effects on the aggregate lending response.

dynamic setup. As we discussed in Section 5, in the light of the Harris, Opp, and Opp (2023) result, changes in capital requirements can have strong effects on the composition of credit. In addition, risk-shifting incentives will also be strongly affected by the state of the economic and financial cycles. For instance, they may be much less strong right after a downturn, when the banking sector has shrunk and competition is effectively muted. In that case, prospective scarcity rents increase banks franchise value (Martinez-Miera and Suarez (2014)). Likewise, bank profitability will be affected by monetary policy (see Drechsler, Savov, and Schnabl 2021; Drechsler et al. 2023) for analyses of the deposit franchise, and (Wang 2018); (Döttling 2023)) on the effect of the zero lower bound). Through the lenses of our static framework, one can interpret such mechanisms as variation in both in the competitive environment, and in the joint distribution of legacy portfolio and investment opportunities. Reinforcing the notion that one should not expect a stable relationship between changes of capital requirements and lending.

References

Acharya, V. V., Bergant, K., Crosignani, M., Eisert, T., & McCann, F. (2022). The anatomy of the transmission of macroprudential policies. *The Journal of Finance, 77*(5), 2533-2575.

Anginer, D., Demirguc-Kunt, A., & Zhu, M. (2014). How does deposit insurance affect bank risk? Evidence from the recent crisis. *Journal of Banking & Finance, 48*, 312-321.

Arbatli-Saxegaard, E. C., & Juelsrud, R. E. (2022). Capital requirement reductions, heterogeneity, and real economic outcomes. *International Journal of Central Banking, 18*(2), 349-401.

Bahaj, S., Bridges, J., Malherbe, F., & O'Neill, C. (2016). What determines how banks respond to changes in capital requirements? *Bank of England Staff Working Paper No. 593*, April.

Bahaj, S., Lattanzio, C., & Malherbe, F. (2024). Within-borrower variation and equilibrium credit supply. *Mimeo*, UCL.

Bahaj, S., & Malherbe, F. (2020). The forced safety effect: How higher capital requirements can increase bank lending. *The Journal of Finance, 75*(6), 3013-3053.

Bahaj, S., & Malherbe, F. (2024). The cross-border effects of bank capital regulation. *Journal of Financial Economics*, *160*, 103912.

Bassett, W. F., & Berrospide, J. M. (2018). The impact of post stress tests capital on bank lending. *Finance and Economics Discussion Series 2018-087*, Board of Governors of the Federal Reserve System (U.S.).

Begenau, J. (2020). Capital requirements, risk choice, and liquidity provision in a business-cycle model. *Journal of Financial Economics, 136*, 355-378.

Behn, M., Haselmann, R., & Wachtel, P. (2016). Procyclical capital regulation and lending. *Journal of Finance, 71*(2), 919-956.

Berg, T., Reisinger, M., & Streitz, D. (2021). Spillover effects in empirical corporate finance. *Journal of Financial Economics, 142*(3), 1109-1127.

Berrospide, J. M., & Edge, R. M. (2024). Bank capital buffers and lending, firm financing and spending: What can we learn from five years of stress test results?. *Journal of Financial Intermediation*, *57*, 101061.

Calomiris, C. W., & Chen, S. (2022). The spread of deposit insurance and the global rise in bank asset risk since the 1970s. *Journal of Financial Intermediation*, *49*, 100881.

Calomiris, C. W., & Jaremski, M. (2019). Stealing deposits: Deposit insurance, risk-taking, and the removal of market discipline in early 20th-century banks. *The Journal of Finance, 74*(2), 711-754.

Chang, B., Gomez, M., & Hong, H. (2023). Sorting out the effect of credit supply. *Journal of Financial Economics*, *150*(3), 103719.

Cucic, D., Iyer, R., Kokas, S., Peydró, J. L., & Pica, S. (2024). Distortive effects of deposit insurance: Administrative evidence from deposit and loan accounts. *Mimeo.*

Danisewicz, P., Lee, C. H., & Schaeck, K. (2022). Private deposit insurance, deposit flows, bank lending, and moral hazard. *Journal of Financial Intermediation*, *52*, 100967.

Dell'Ariccia, G., & Marquez, R. (2006). Competition among regulators and credit market integration. *Journal of Financial Economics, 79*(2), 401-430.

DeMarzo, P., Krishnamurthy, A., & Nagel, S. (2024). Interest rate risk in banking. *Mimeo*.

Döttling, R. (2024). Bank capital regulation in a zero interest environment. *Journal of the European Economic Association*.

Drechsler, I., Savov, A., & Schnabl, P. (2021). Banking on deposits: Maturity transformation without interest rate risk. *The Journal of Finance, 76*(3), 1091-1143.

Drechsler, I., Savov, A., Schnabl, P., & Wang, O. (2023). Banking on uninsured deposits. *National Bureau of Economic Research*.

Du, W., Tepper, A., & Verdelhan, A. (2018). Deviations from covered interest rate parity. *Journal of Finance, 73*(3), 915-957.

Duffie, D., Andersen, L., & Song, Y. (2019). Funding value adjustments. *Journal of Finance, 74*(2), 145-192.

Favara, G., Ivanov, I., & Rezende, M. (2021). GSIB surcharges and bank lending: Evidence from US corporate loan data. *Journal of Financial Economics, 142*(3), 1426-1443.

Fraisse, H., Le, M., & Thesmar, D. (2019). The real effects of bank capital requirements. *Management Science, 66*(1), 5-23.

Gropp, R., Mosk, T., Ongena, S., & Wix, C. (2019). Banks response to higher capital requirements: Evidence from a quasi-natural experiment. *The Review of Financial Studies, 32*(1), 266-299.

Hanson, S. G., Malkhozov, A., & Venter, G. (2024). Demand-and-supply imbalance risk and long-term swap spreads. *Journal of Financial Economics*, *154*, 103814.

Harris, M., Opp, C. C., & Opp, M. M. (2024). Intermediary capital and the credit market. *Management Science*.

Haubrich, J. G. (2020). A brief history of bank capital requirements in the United States. *Economic Commentary, 2020-05*.

Hellmann, T. F., Murdock, K. C., & Stiglitz, J. E. (2000). Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough? *American Economic Review, 90*(1), 147-165.

Huber, K. (2023). Estimating general equilibrium spillovers of large-scale shocks. *The Review of Financial Studies, 36*(4), 1548-1584.

Imbierowicz, B., Kragh, J., & Rangvid, J. (2018). Time-varying capital requirements and disclosure rules: Effects on capitalization and lending decisions. *Journal of Money, Credit and Banking, 50*(4), 573-602.

Jiménez, G., Ongena, S., Peydro, J. L., & Saurina, J. (2017). Macroprudential policy, countercyclical bank capital buffers, and credit supply: Evidence from the Spanish dynamic provisioning experiments. *Journal of Political Economy, 125*(6), 2126-2177.

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

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

Keeley, M. C. (1990). Deposit insurance, risk, and market power in banking. *The American Economic Review*, 1183-1200.

Khwaja, A. I., & Mian, A. (2008). Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *American Economic Review, 98*(4), 1413-1442.

Kreps, D. M., & Scheinkman, J. A. (1983). Quantity precommitment and Bertrand competition yield Cournot outcomes. *The Bell Journal of Economics, 14*(2), 326-337.

Laeven, L., & Levine, R. (2009). Bank governance, regulation and risk taking. *Journal of Financial Economics, 93*(2), 259-275.

Landier, A., Sraer, D. A., & Thesmar, D. (2015). The risk-shifting hypothesis: Evidence from subprime originations. In *AFA 2012 Chicago Meetings Paper*.

Lattanzio, C. (2024). How to model bank competition: The case for Cournot. *Mimeo*.

Malherbe, F. (2020). Optimal capital requirements over the business and financial cycles. *American Economic Journal: Macroeconomics, 12*(3), 139-174.

Malherbe, F., & McMahon, M. (2024). Beyond Pangloss: Financial sector origins of inefficient economic booms. *Journal of Monetary Economics*, 103558.

Martinez-Miera, D., & Suarez, J. (2014). A macroeconomic model of endogenous systemic risk taking. *Mimeo*, CEMFI.

Merton, R. C. (1977). An analytic derivation of the cost of deposit insurance and loan guarantees an application of modern option pricing theory. *Journal of Banking & Finance*, *1*(1), 3-11.

Mian, A., Sarto, A., & Sufi, A. (2023). Estimating credit multipliers. *Working Paper*.

Oehmke, M., & Opp, M. M. (2023). Green capital requirements. *Mimeo*, LSE.

Paravisini, D., Rappoport, V., & Schnabl, P. (2023). Specialization in bank lending: Evidence from exporting firms. *The Journal of Finance, 78*(4), 2049-2085.

Puri, M., Rocholl, J., & Steffen, S. (2011). Global retail lending in the aftermath of the US financial crisis: Distinguishing between supply and demand effects. *Journal of Financial Economics, 100*(3), 556-578.

Schliephake, E., & Kirstein, R. (2013). Strategic effects of regulatory capital requirements in imperfect banking competition. *Journal of Money, Credit and Banking, 45*(4), 675-700.

Thakor, A. V. (1996). Capital requirements, monetary policy, and aggregate bank lending: Theory and empirical evidence. *The Journal of Finance, 51*(1), 279-324.

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

Wang, O. (2024). Banks, low interest rates, and monetary policy transmission. *Journal of Finance, forthcoming.*