Volume 7, Number 4 December 2011

IJCB
INTERNATIONAL JOURNAL OF CENTRAL BANKING
Volume 7, Number 4
December 2011

IJCB

INTERNATIONAL JOURNAL OF CENTRAL BANKING

Special Issue: The Real and Financial Effects of Basel III

IJCB International Journal of Central Banking

Volume 7, Number 4 December 2011

Special Issue: The Real and Financial Effects of Basel III

The *International Journal of Central Banking* is published quarterly (ISSN: 1815-4654). Online access to the publication is available free of charge at **www.ijcb.org**. Individual print subscriptions are available at an annual rate of $100 (USD).

Print subscription orders may be placed by phone (001 415 974 2035), via fax (001 415 974 2168), or by e-mail (editor@ijcb.org).

Renewals, claims, address changes, and requests for permission to reprint material from this journal should be addressed to:

International Journal of Central Banking
Economic Research Department
Federal Reserve Bank of San Francisco
101 Market Street
San Francisco, CA 94105
USA

Phone: 001 415 974 2035
Fax: 001 415 974 2168
E-mail: editor@ijcb.org

ISSN: 1815-4654

The Real and Financial Effects of Basel III

Introduction to a Special Issue of the International Journal of Central Banking

Douglas Gale, Rafael Repullo, and Frank Smets

Following the widespread financial instability of recent years, the Basel Committee on Banking Supervision has put forward reform proposals, commonly referred to as Basel III, to increase the resilience of the banking sector. The reform package is a major overhaul of the current regulatory framework, as it includes a comprehensive set of rules encompassing tighter definitions of capital, improved risk capture, a non-risk-based leverage ratio, a framework for capital conservation and countercyclical buffers, and a novel regime for liquidity risk. In addition, work on the regulatory response to the risks posed by systemically important financial institutions is under way. This issue of the *International Journal of Central Banking* contains papers, discussions, and commentaries that address some of the real and financial effects of Basel III. The papers were presented at the third IJCB Financial Stability Conference hosted by the Bank of England on May 26–27, 2011.

The first paper, "A Pigovian Approach to Liquidity Regulation," deals with liquidity regulation. Enrico Perotti and Javier Suarez use a model in which short-term funding enables credit growth but generates negative systemic risk externalities. They show that a Pigovian tax on short-term funding is optimal when banks differ in credit opportunities. In contrast, when banks differ mostly in gambling incentives, excess credit and liquidity risk are best controlled with a quantity constraint such as a net stable funding requirement. More generally, an optimal policy involves both types of tools. The paper is discussed by Ernst-Ludwig von Thadden.

The second paper—"Macroeconomic Propagation under Different Regulatory Regimes: Evidence from an Estimated DSGE Model for the Euro Area," by Matthieu Darracq Pariès, Christoffer Kok Sørensen and Diego Rodriguez-Palenzuela—examines the economic implications of increasing capital requirements using an estimated euro-area DSGE model with financially constrained households and firms and an oligopolistic banking sector facing capital constraints. The authors show, among other results, that the introduction of more stringent capital requirements (as proposed under Basel III) leads to a transitory negative impact on output, although gradual implementation of the new regulation may smooth out the transitional costs to the economy. The paper is discussed by Andrew Powell.

Some of the implications of capital regulation are also discussed in the third paper, “Capital Regulation and Tail Risk,” by Enrico Perotti, Lev Ratnovski, and Razvan Vlahu. These authors analyze how the presence of tail risk affects the relationship between bank capital and bank risk taking. They show that higher capital requirements may have an unintended effect of enabling banks to take more tail risk without the fear of breaching the minimal capital ratio in non-tail risk states. The paper is discussed by Andrew Winton.

The commentary by Richard Herring, “The Capital Conundrum,” relates to the paper by Anat Admati, Peter DeMarzo, Martin Hellwig, and Paul Pfleiderer entitled “Fallacies, Irrelevant Facts, and Myths in the Discussion on Capital Regulation: Why Bank Equity Is *Not* Expensive,” which was presented at the conference but is not published in this issue. He argues that there is a strong case for requiring much higher equity-to-asset ratios and restating these regulatory requirements in terms of proxies for market values rather than relying exclusively on accounting measures.

The last two papers deal with the framework for countercyclical capital buffers. In “Anchoring Countercyclical Capital Buffers: The Role of Credit Aggregates,” Mathias Drehmann, Claudio Borio, and Kostas Tsatsaronis investigate the performance of different variables as anchors for setting the level of the countercyclical capital requirements for banks and conclude that a real-time measure of the credit-to-GDP gap is one of the best predictors of the build-up of systemwide vulnerabilities and of the resulting banking crises. In his discussion, Òscar Jordà proposes a formal way of choosing a threshold for this indicator so as to balance the costs and benefits of each alternative.

Finally, in “The Unreliability of Credit-to-GDP Ratio Gaps in Real Time: Implications for Countercyclical Capital Buffers,” Rochelle Edge and Ralf Meisenzahl emphasize that ex post revisions of the credit-to-GDP gap in the United States are sizable and often as large as the gap itself and consider the potential costs of this mismeasurement. Simon van Norden discusses the paper and argues that the evidence of ex post revisions is not necessarily relevant for the ex ante use of the indicator.

A Pigovian Approach to Liquidity Regulation*

Enrico Perotti[a] and Javier Suarez[b]
[a]University of Amsterdam, De Nederlandsche Bank,
and Duisenberg School of Finance
[b]CEMFI and CEPR

This paper discusses liquidity regulation when short-term funding enables credit growth but generates negative systemic risk externalities. It focuses on the relative merit of price versus quantity rules, showing how they target different incentives for risk creation.

When banks differ in credit opportunities, a Pigovian tax on short-term funding is efficient in containing risk and preserving credit quality, while quantity-based funding ratios are distortionary. Liquidity buffers are either fully ineffective or similar to a Pigovian tax with deadweight costs. Critically, they may be least binding when excess credit incentives are strongest.

When banks differ instead mostly in gambling incentives (due to low charter value or overconfidence), excess credit and liquidity risk are best controlled with net funding ratios. Taxes on short-term funding emerge again as efficient when capital or liquidity ratios keep risk-shifting incentives under control. In general, an optimal policy should involve both types of tools.

JEL Codes: G21, G28.

1. Introduction

The recent crisis has provided a clear rationale for the regulation of banks' refinancing risk, a critical gap in the Basel II framework. This

*We have greatly benefited from numerous discussions with academics and policymakers on our policy writings on the regulation of liquidity and on this paper. Our special thanks to Viral Acharya, Max Bruche, Willem Buiter, Oliver Burkart, Douglas Gale, Charles Goodhart, Nigel Jenkinson, Laura Kodres, Arvind Krishnamurthy, David Martinez-Miera, Rafael Repullo, Jeremy Stein, and Ernst-Ludwig von Thadden for insightful suggestions and comments. Contact e-mails: e.c.perotti@uva.nl, suarez@cemfi.es.

paper studies the effectiveness of different approaches to liquidity regulation.

The basic trade-off of short-term funding is that rapid expansion of credit may only be funded by attracting short-term funding (for instance, because deposit supply can be expanded only slowly, or because short-term market lenders do not need to be very informed about new credit choices), but this creates refinancing risk, especially in situations where there are doubts about banks' fundamental solvency or evidence of excessive risk taking on the asset side as well.[1] Sudden withdrawals may lead to disruptive liquidity runs (Diamond and Dybvig 1983), and cause fire sales or counterparty risk externalities which affect other intermediaries exposed to short-term funding (Brunnermeier 2009; Allen, Babus, and Carletti 2010). As a result, each bank's funding decision has an impact on the vulnerability of other banks to liquidity risk, causing a negative externality. Even if an individual bank's funding decision takes into account its own exposure to refinancing risk, it will not internalize its systemwide effect (Perotti and Suarez 2009). Because of the wedge between the net private value of short-term funding and its social cost, banks will rely excessively on short-term funding. A prime example is the massive build-up in wholesale funding which supported the recent securitization wave, and the overnight credit (repo) growth during 2002–07, which grew explosively to a volume over ten trillion dollars (Gorton 2009).

In the tradition of externality regulation led by Weitzman (1974), we assess the performance of Pigovian taxes (aimed at equating private and social liquidity costs) and quantity regulations in containing this systemic externality. As in Weitzman, the optimal regulatory tool depends on the response elasticity of banks, recognizing that the regulator is informationally constrained in targeting individual bank characteristics.[2] Our results show how the industry response

[1] In this paper, we adopt the same strategy as the Basel Committee on Banking Supervision (BCBS) and, for simplicity, discuss liquidity risk as if it were distinct and separable from other sources of risk. Arguably, liquidity risk and solvency risk are hardly separable, both in practice, as it is hard to tell apart insolvent banks from those with pure liquidity problems, and conceptually, since solvency concerns are at the root of most liquidity crises.

[2] Our analysis is also related to the classical discussion by Poole (1970) on the optimality of price or quantity monetary policy instruments when the system to regulate is affected by several types of shocks. For an earlier discussion on price vs. quantity regulation in banking, see Keeton (1990).

to regulation depends on the composition of bank characteristics. The model recognizes that banks differ in their credit ability and their incentives to take risk. Banks earn decreasing returns to expand credit to their (monitored) borrowers, so better banks naturally lend more. Shareholders of less-capitalized banks gain from investing in poor gambles, since they retain the upside and shift downside risk to the public safety net.[3] Depending on the dominant source of heterogeneity, the socially efficient solution may be attained with Pigovian taxes, quantity regulations, or a combination of both.

To facilitate the discussion, we first analyze the impact of regulation under either investment-quality or risk-taking heterogeneity.[4] When banks differ only in capacity to lend profitably (reflecting credit assessment capability or access to credit opportunities), a simple flat-rate Pigovian tax on short-term funding (possibly scaled up by the systemic importance of each bank) implements the efficient allocation. The intuition is that liquidity risk levies allow better banks to lend more, without requiring the regulators to be able to identify them. In this context, quantity-based instruments such as the net stable funding ratio or the liquidity coverage ratio introduced by Basel III (see Basel Committee on Banking Supervision 2010) may improve over the unregulated equilibrium but are generally distortionary. An optimal quantity-based approach would require introducing contingency on individual bank characteristics, most of which are imprecisely measured or unobservable.[5]

More precisely, net stable funding ratios which impose an upper limit on short-term debt do reduce overall liquidity risk but redistribute liquidity risk inefficiently across banks. Banks with better credit opportunities will be constrained, while the reduced systemic

[3]An alternative view of gambling incentives is that they are driven by self-interested and overconfident managers, which view excessive risks as profitable.

[4]Each form of heterogeneity leads to a situation akin to each of the polar cases that Weitzman (1974, p. 485) describes in terms of the "curvature" of the social benefit function and the private cost function relevant to his analysis—he finds that price (quantity) regulation dominates when the social benefit (private cost) function is linear.

[5]Quantity requirements may be indexed to measures of the "systemic importance" of each bank such as those recently considered in regulatory discussions (size, interconnectedness, lack of substitutability, global cross-jurisdictional activity, and complexity), but it is much harder to find good proxies for banks' credit opportunities.

risk actually encourages banks with low credit ability (for which the requirement is not binding) to expand.

Liquidity coverage ratios which require banks to hold fractional reserves of liquid assets against short-term funding work as a de facto tax but one that can turn out to be ineffective or imply large deadweight losses.[6] When the yield on liquid assets equals the cost of short-term liabilities (roughly the case in normal times, and certainly prior to the crisis), buffers impose no net cost to stacking liquidity. Banks will simply increase their gross short-term funding to keep their "net" short-term funding (i.e., minus the buffers) as high as in the unregulated equilibrium. The only effect is an artificial demand for liquid assets—traditionally kept in money market mutual funds rather than banks—that might be redirected to banks following the new requirement.

When the spread between bank borrowing costs and liquid asset yields is positive, a liquidity requirement operates as a tax on short-term funding that involves deadweight costs and has its effective tax rate partly determined by market spreads.[7] In the recent experience, the interbank spread over safe assets was minimal just as aggregate liquidity risk was building up, and it jumped up with the start of the crisis. This means that liquidity requirements would have to be increased in good times and reduced in bad times so as to avoid making them a source of further banking system procyclicality.

Studying variation in risk-shifting incentives (correlated with charter value or another determinant of risk-taking tendencies, such as overconfidence) alters the results radically. Low-charter-value (or more risk-loving) banks have strong incentives to gamble to shift risk to the deposit insurance provider (Keeley 1990). We show that decisions driven by such gambling incentives are not properly deterred by levies, while quantity constraints are more effective. Both short-term funding limits (e.g., a net stable funding ratio) and capital requirements can contain risk shifting by limiting the scale of lending. Levies will not be as effective as in the previous scenario because the most gambling-inclined banks will also be the most inclined to

[6]Liquid assets which can be sold at no fire-sale loss in a crisis are essentially cash, central bank reserves, and treasury bills.

[7]Specifically, the tax rate will equal the product of the buffer requirement per unit of short-term funding times the interest rate spread.

pay the tax and expand their risky lending. In this case, quantity instruments such as net funding or capital ratios are best to contain excess credit expansion.

Our analysis identifies the relative merits of price versus quantity instruments and suggests that combining them may be adequate for the simultaneous control of gambling incentives and systemic risk externalities. However, this presumes that the regulator controls only instruments connected to liquidity risk. If strengthening capital requirements is an effective strategy for the control of gambling incentives (Hellmann, Murdock, and Stiglitz 2000; Repullo 2004), the case for levies on short-term funding is considerably reinforced.[8]

Other considerations may qualify the recommendation for the use of price-based or quantity-based instruments. For instance, levies may be less costly to adjust than ratios. First, they might be easier to change for institutional reasons (e.g., if regulatory ratios are embedded in some law or international agreement while the levies are, at least partly, under control of a macroprudential authority). More importantly, they may imply lower adjustment costs at the bank level than changing bank funding volumes on short notice. Similarly, changes in levies are less likely to induce procyclicality, since the Pigovian "tax rate" is directly controlled by the regulator rather than implicitly set by the interaction of some (controlled) quantitative requirement and the (freely fluctuating) market price of the required resource (namely, capital, liquid assets, or stable funding). For preventive policy, controlling time-varying liquidity risk may then be best achieved by a combination of stable ratios and variable levies.

The rest of the paper is organized as follows. Section 2 describes some related literature and some recent evidence on liquidity risk. Section 3 describes the baseline model. Section 4 characterizes the unregulated equilibrium. Section 5 finds the socially optimal allocation. In section 6, we discuss the possibility of restoring efficiency with a Pigovian tax on short-term funding. Section 7 considers alternative quantity-based regulations. In section 8 we analyze the

[8]Hellmann, Murdock, and Stiglitz (2000) point out that "capital requirements also have a perverse effect of harming banks' franchise values, thus encouraging gambling" but the analysis in Repullo (2004) suggests that the standard deleveraging effect tends to dominate the charter value effect.

implications of introducing gambling incentives as a second dimension of bank heterogeneity. Section 9 discusses further implications and extensions of the analysis. Section 10 concludes the paper.

2. Evidence from the Crisis and Related Research

The crisis of 2007–08 has been described as a wholesale bank crisis, or a repo run crisis (Gorton 2009). The rapid withdrawing of short-term debt was responsible for propagation of shocks across investors and markets (Brunnermeier 2009). Brunnermeier and Oemhke (forthcoming) show that creditors have an incentive to shorten their loan maturity, so as to pull out in bad times before other creditors can. This, in turn, causes a lender race to shorten maturity, leading to excessively short-term financing. The consequences are formalized in Martin, Skeie, and von Thadden (2010), where increased collective reliance on repo funding weakens solvency constraints and produces repo runs. Acharya and Viswanathan (2011) model the sudden drying up of liquidity when banks need to refinance short-term debt in bad times. As low asset prices increase incentives for risk shifting, investors may rationally refuse refinancing to illiquid banks.

Acharya, Gale, and Yorulmazer (2011) show that high rollover frequency can reduce the collateral value of risky securities, but they treat debt maturity as exogenous and do not look at the normative implications. Papers emphasizing the possibility of socially inefficient levels of maturity transformation include Huang and Ratnovski (2011), who focus on the deterioration of information production incentives among banks; Farhi and Tirole (2010), where the distortion comes from the expectation of a bail-out; and Segura and Suarez (2011), where the pricing of refinancing during crisis interacts with banks' financial constraints and gives rise to pecuniary externalities linked to banks' funding maturity decisions.

While the role of liquidity risk in the crisis has been evident from the beginning, more precise empirical evidence is now emerging. Acharya and Merrouche (2010) show that UK banks with more wholesale funding and fire-sale losses in 2007–08 contributed more to the transmission of shocks to the interbank market. A concrete measure of the role of short-term debt played in the credit boom, and its demise comes from the explosive rise of repo (overnight) financing in the last years and its rapid deflation since the panic (Gorton 2009).

Repo funding evaporated in the crisis, leading to bursts of front running in the sales of repossessed securities. Adrian and Brunnermeier (2008) present evidence of the correlation between banks' use of short-term wholesale funding and their proposed measure of banks' contribution to systemic risk (CoVaR). A similar result emerges in Acharya et al. (2010).

The leading causes of external effects from refinancing risk have been identified as losses due to fire sales and collective fears about counterparty risk amplified by simultaneous refinancing choices. They have motivated proposals on the creation of private or public clearing arrangements to limit the effects of runs, though purely private arrangements are not expected to be sufficient in systemic liquidity runs. Acharya and Öncü (2010) argue for the establishment of a Repo Resolution Authority to take over repo positions in a systemic event, paying out a fraction of their claims and liquidating the collateral in an orderly fashion. This would force investors to bear any residual loss. On the opposite front, Gorton (2009) has proposed stopping fire sales of seized collateral by a blanket state guarantee, while Gorton and Metrick (2010) propose creating special vehicles they call narrow banks to hold such assets, backed by a public guarantee.

Another critical issue is the consequences of ex post liquidity bailouts (Farhi and Tirole 2010). The expectation that in a systemic run there is no choice but to provide liquidity to mismatched intermediaries may produce an ex ante moral hazard problem whereby individual institutions take even fewer precautions against crises. This highlights the urgency of measures to contain the private creation of liquidity risk. Finally, systemic crises are the source of important fiscal and real losses not fully internalized by those who make the decisions that lead to the accumulation of systemic liquidity risk (Laeven and Valencia 2010), making a clear case for regulation.

The paper is related to several other strands of the academic literature which would take too long to revise in a systematic manner. These include the corporate finance and banking literatures on the potentially beneficial incentive effects of short-term funding (e.g., Calomiris and Kahn 1991, Diamond and Rajan 2001, and Huberman and Repullo 2010), on the connection between short-term funding and banks' vulnerability to panics and contagion (e.g., Allen and Gale 2000, Rochet and Vives 2004, and Allen, Babus, and Carletti

2010), and on externalities related to other financial decisions, such as diversification decisions (Wagner 2010) or decisions regarding the supply of credit over the business cycle (Lorenzoni 2008; Jeanne and Korinek 2010). Finally, our analysis is also connected to a vast economic literature about the choice between quantity-based and price-based regulation in specific setups.[9]

3. The Model

Consider a one-period model of a banking economy in which all agents are risk neutral. The banking system is made up of a continuum of heterogenous banks run by their owners with the objective of maximizing their expected net present value (NPV). To start with, we assume that banks differ in a parameter θ that affects the NPV that they can generate using short-term funding, whose amount will be their only decision variable for the time being.[10] The parameter θ follows a continuous distribution with positive density $f(\theta)$ over the interval $[0,1]$. Assuming w.l.o.g. that all banks of each class θ behave symmetrically, the short-term funding decision of each bank of class θ is denoted by $x(\theta) \in [0,\infty)$.

We postulate that the expected NPV associated with a decision x by a bank of class θ can be written as

$$v(x, X, \theta) = \pi(x, \theta) - \varepsilon(x, \theta)c(X), \qquad (1)$$

where X is a measure of the aggregate systemic risk implied by the individual funding decisions of all banks, $\pi(x, \theta)$ is the NPV generated if no systemic liquidity crisis occurs, and $\varepsilon(x, \theta)c(X)$ is the expected NPV loss due to the possibility of a systemic liquidity crisis. To facilitate the presentation, we assume a multiplicative decomposition of the expected crisis losses in two terms: the term $\varepsilon(x, \theta) \geq 0$, which captures the purely individual contribution of the funding decision x and the individual characteristic θ to the vulnerability of the bank, and the term $c(X) \geq 0$, which captures the influence of other banks' funding decisions on systemic crisis costs.

[9]See contributions such as Glaeser and Shleifer (2001) and Kaplow and Shavell (2002) for an overview of the literature.

[10]In section 8, we introduce a second dimension of bank heterogeneity directed to capture differences in banks' gambling incentives.

We assume that $\pi(x, \theta)$ is increasing and differentiable in its two arguments, strictly concave in x, and with a positive cross-derivative, $\pi_{x\theta} > 0$, so that a larger θ implies a larger capability to extract value from short-term funding. To facilitate obtaining interior solutions in x and monotone comparative statics with respect to θ, we also assume that $\varepsilon(x, \theta)$ is differentiable, increasing and weakly convex in x, non-increasing in θ, and with $\varepsilon_{x\theta} \leq 0$. Finally, we assume $c(X)$ to be increasing, differentiable, and weakly convex in X.

A structural story consistent with this specification might be that $\pi(x, \theta)$ captures the profitability, in the absence of a systemic liquidity crisis, of using short-term funding to expand lending, $\varepsilon(x, \theta)$ captures the probability that the bank faces refinancing problems in a liquidity crisis and has to accommodate them by, say, selling its assets, and $c(X)$ denotes the net liquidation losses incurred in such an event. Notice that $c(X)$ might be increasing in X due to the impact on liquidation values of concurrent sales from troubled banks (e.g., under some cash-in-the-market pricing logic or simply because the alternative users of the liquidated assets face marginally decreasing returns).[11] Here θ can be taken as a measure of a bank's credit ability or any other determinant of the marginal net value of its investments.

The key results below would be robust to essentially any specification of the aggregator $X = g(\{x(\theta)\})$, where $\{x(\theta)\}$ is the schedule of the short-term funding used by the banks in each class $\theta \in [0, 1]$ and we have $\partial g/\partial x(\theta) \geq 0$ for all θ. For concreteness, however, we focus on the case in which aggregate systemic liquidity risk can be measured as the simple sum of all individual decisions:

$$X = g(\{x(\theta)\}) = \int_0^1 x(\theta) f(\theta) d\theta. \qquad (2)$$

In section 9, we will discuss how to adapt our main results to the case in which banks also differ in a "systemic importance" factor

[11]Of course, an increasing $c(X)$ may also partly reflect that X increases the very probability of a systemic crisis. For example, the more vulnerable banks' funding structures are, the more likely it is that asset-side shocks such as a housing market bust or a stock market crash get transformed into a systemic liquidity shock.

that affects the weight of the contribution of their short-term funding to X.

We assume that all investors, except bank owners, have the opportunity to invest their wealth at exogenously given market rates and provide funding at competitive terms, hence obtaining a zero NPV from dealing with the banks. Additionally, the model attributes the whole NPV associated with bank investment opportunities to the bank owners, which means that the overall value of the banks to their owners is the natural measure of social welfare W in this economy. To properly interpret this measure, notice that, in reality, banks' investments consist of lending to (or investing in assets issued by) other agents and, in a competitive environment, a significant part of the NPV that constitutes our welfare measure will tend to pass to the borrowers (or the issuers of the assets) through endogenous improvements in lending conditions (or increases in asset prices). We attribute all the gains to the bank owners simply because we abstract from modeling the details of such a pass-through process. Formally, our expression for social welfare is

$$\begin{aligned} W(\{x(\theta)\}) &= \int_0^1 v(x(\theta), X, \theta) f(\theta) d\theta \\ &= \int_0^1 [\pi(x(\theta), \theta) - \varepsilon(x(\theta), \theta) c(X)] f(\theta) d\theta. \qquad (3) \end{aligned}$$

Notice that the short-term funding decision x of any bank of class θ determines, via $\varepsilon(x, \theta)$, the vulnerability of that very bank to a systemic crisis and also, via $c(X)$, the likelihood and/or costs of a systemic crisis to all other banks.

4. Equilibrium

In an unregulated competitive equilibrium, each bank chooses x so as to maximize its own expected NPV, $v(x, X, \theta)$, taking X as given. So an *unregulated competitive equilibrium* is a pair $(\{x^e(\theta)\}, X^e)$ that satisfies

(i) $x^e(\theta) = \arg\max_x \{\pi(x, \theta) - \varepsilon(x, \theta) c(X^e)\}$ for all $\theta \in [0, 1]$,
(ii) $X^e = \int_0^1 x^e(\theta) f(\theta) d\theta$.

Taking into account the implicit non-negativity constraint on x, banks' privately optimal choice of x under given values of θ and X can be described as the maximum between zero (which corresponds to a corner solution) and the unique solution $y(\theta, X)$ to the first-order condition

$$\pi_x(y(\theta, X), \theta) - \varepsilon_x(y(\theta, X), \theta)c(X) = 0, \tag{4}$$

which characterizes an interior solution. Given the assumed properties of the functions involved in this first-order condition, the implicit function theorem implies that $y(\theta, X)$ is increasing in θ and decreasing in X.[12]

The equilibrium value of X can be found as the fixed point of the auxiliary function $h(X) = \int_0^1 \max\{0, y(\theta, X)\} f(\theta) d\theta$. This function is decreasing in X insofar as $y(\theta, X) > 0$ for some positive measure set of values of θ. If we assume $y(1, 0) > 0$, so that at least the banks with the largest valuation of short-term funding (those with $\theta = 1$) find it worthy to use some of it if $X = 0$, then it is guaranteed that $h(0) > 0$ and, by standard arguments, there exists a unique fixed point $X^e > 0$.

In what follows we will refer to an equilibrium as *interior* if it satisfies $x^e(\theta) > 0$ for all $\theta > 0$, in which case the first-order condition

$$\pi_x(x^e(\theta), \theta) - \varepsilon_x(x^e(\theta), \theta)c(X^e) = 0, \tag{5}$$

with $X^e = \int_0^1 x^e(\theta) f(\theta) d\theta$, is satisfied for all $\theta \in [0, 1]$. Guaranteeing that the equilibrium is interior requires having

$$\pi_x(0, 0) - \varepsilon_x(0, 0)c(X) \geq 0 \tag{6}$$

for a sufficiently large X (e.g., larger than the possibly emerging X^e), so that even the banks with the lowest inclination for short-term funding (those with $\theta = 0$) want to use some of it in equilibrium.[13]

[12] Recall that we have assumed $\pi_{x\theta} > 0$ and $\varepsilon_{x\theta} \leq 0$.

[13] A sufficient condition for this would be to have $\pi_x(0, 0) - \varepsilon_x(0, 0)c(\overline{X}) \geq 0$ with $\overline{X}$ implicitly defined by the equation $\pi_x(\overline{X}, 1) - \varepsilon_x(\overline{X}, 1)c(\overline{X}) = 0$, which would characterize the equilibrium value of aggregate systemic risk in the hypothetical scenario in which all banks were of the highest type $\theta = 1$. To obtain most of the results below, we need not constrain attention to interior equilibria, but a full discussion of the solutions involving $x(\theta) = 0$ for some θ would make the presentation unnecessarily cumbersome.

As shown below, the presence of systemic risk externalities will make the conditions stated in (5) incompatible with social efficiency.

5. The Social Planners' Problem

The socially optimal allocation of short-term funding across banks can be found be maximizing social welfare W, taking into account the influence of each individual bank funding strategy on X. Formally, a *socially optimal allocation* can be defined as a pair $(\{x^*(\theta)\}, X^*)$ that satisfies

$$\begin{aligned}
&(\{x^*(\theta)\}, X^*) \\
&= \arg \max_{(\{x(\theta)\},X)} \int_0^1 [\pi(x(\theta),\theta) - \varepsilon(x(\theta),\theta)c(X^*)]f(\theta)d\theta \\
&\quad\quad \text{s.t.:} \quad \int_0^1 x(\theta)f(\theta)d\theta = X^*.
\end{aligned} \tag{7}$$

After substituting the constraint in the objective function, we can find the social optimum by solving the system of equations

$$\begin{aligned}
\{x^*(\theta)\} = \arg \max_{\{x(\theta)\}} \int_0^1 \Bigg[&\pi(x(\theta),\theta) \\
&- \varepsilon(x(\theta),\theta)c\left(\int_0^1 x(z)f(z)dz\right)\Bigg] f(\theta)d\theta
\end{aligned} \tag{8}$$

for $\theta \in [0, 1]$, and then finding $X^* = \int_0^1 x^*(\theta)f(\theta)d\theta$, recursively.

The assumptions adopted in section 3 guarantee the existence of a unique socially optimal allocation. For values of θ with $x^*(\theta) > 0$, the first-order condition associated with the maximization in (8) establishes

$$\pi_x(x^*(\theta),\theta) - \varepsilon_x(x^*(\theta),\theta)c(X^*) - E_z(\varepsilon(x^*(z),z))c'(X^*) = 0, \tag{9}$$

where $E_z(\varepsilon(x^*(z),z)) = \int_0^1 \varepsilon(x^*(z),z)f(z)dz$. Intuitively, and similarly to the equilibrium allocation, insofar as $x^*(\theta) > 0$, the socially optimal allocation will assign larger short-term funding to the banks with higher θ. Hence, the socially optimal allocation will be interior if and only if

$$\pi_x(0,0) - \varepsilon_x(0,0)c(X^*) - E_\theta(\varepsilon(x^*(\theta),\theta))c'(X^*) > 0. \tag{10}$$

An interior socially optimal allocation can then be guaranteed if, e.g., the profit function satisfies $\pi_x(0,0) \to \infty$ and the functions $\varepsilon(x,\theta)$ and $c(X)$ have finite derivatives with respect to x and X, respectively.[14]

Relative to the conditions for individual bank optimization in an interior equilibrium, given in (5), the conditions in (9) add a third, negative term that reflects the marginal external cost associated with each $x(\theta)$. The marginal external cost relevant for a bank of class θ is made of two multiplicative factors: the average vulnerability of all the banks in the system to a systemic crisis, $E_z(\varepsilon(x^*(z),z))$, and the marginal effect of aggregate funding risk on systemic crisis costs, $c'(X^*)$.

The presence of the external cost term in (9) implies that any equilibrium allocation $\{x^e(\theta)\}$ with $x^e(\theta) > 0$ for some positive measure set of θs cannot coincide with the socially optimal allocation $\{x^*(\theta)\}$. To see this, notice that if $\{x^e(\theta)\}$ were socially optimal, then, by (9), for any θ with $x^e(\theta) > 0$, we should have

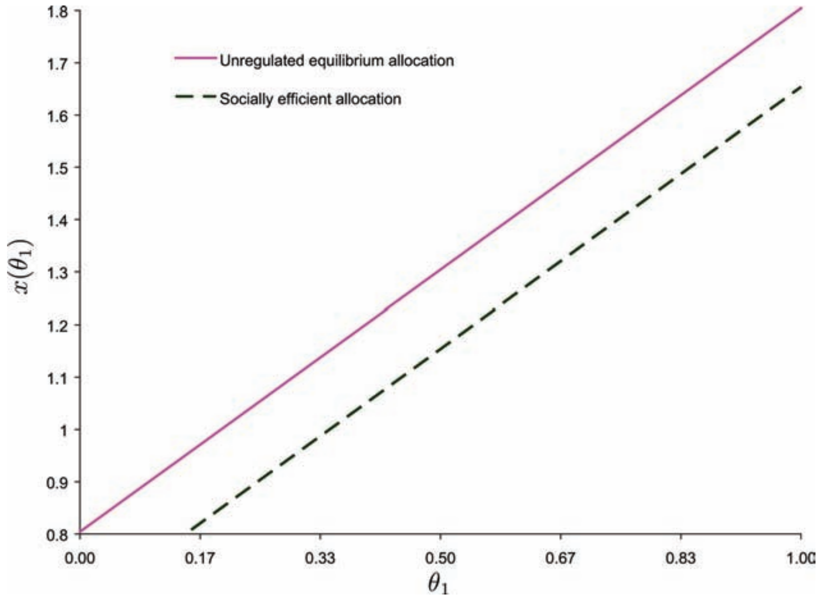
$$\pi_x(x^e(\theta),\theta) - \varepsilon_x(x^e(\theta),\theta)c(X^e) - E_z(\varepsilon(x^e(z),z))c'(X^e) = 0. \quad (11)$$

However, the first two terms add up to zero, by (5), while the third term is strictly negative since $E_z(\varepsilon(x^e(z),z)) > 0$ and $c'(X^e) > 0$. So (11) does not hold.

The presence of a negative externality associated with the contribution of banks' short-term funding to aggregate systemic risk suggests that the socially optimal allocation will generally involve a strictly lower level of systemic risk than the equilibrium allocation. Indeed, under the hypothesis that $X^* \geq X^e > 0$, (5) and (9) would imply that $x^e(\theta) > x^*(\theta)$ for all θ with $x^e(\theta) > 0$ (and $x^e(\theta) = x^*(\theta) = 0$ otherwise), which would obviously mean $X^e > X^*$ and, hence, would contradict the hypothesis. In practical terms, the results imply that in the unregulated competitive

[14]In the same spirit as in the condition given in footnote 4, an alternative sufficient condition would be to have $\pi_x(0,0) - \varepsilon_x(0,0)c(\overline{X}) - \varepsilon(\overline{X},0)c'(\overline{X}) \geq 0$ with $\overline{X}$ implicitly defined by the equation $\pi_x(\overline{X},1) - \varepsilon_x(\overline{X},1)c(\overline{X}) - \varepsilon(\overline{X},1)c'(\overline{X}) = 0$, which would characterize the socially optimal level of aggregate systemic risk in the hypothetical scenario in which all banks were of the highest type $\theta = 1$. The inequality above uses $\varepsilon(\overline{X},0)$ as an upper bound to $E_\theta(\varepsilon(x^*(\theta),\theta))$, since the vulnerability function $\varepsilon(x,\theta)$ is increasing in x and decreasing in θ.

Figure 1. Unregulated Equilibrium vs. Socially Optimal Allocations



equilibrium, sufficiently many banks (if not all) use more short-term funding than they would use under the socially optimal allocation.[15] The key insights from this section are summarized in the following proposition.

PROPOSITION 1. *The presence of systemic externalities associated with banks' funding decisions,* $c'(X) > 0$, *makes the equilibrium allocation socially inefficient and characterized by an excessive aggregate systemic risk* $X^e > X^*$.

Figure 1 depicts the short-term funding decisions that correspond to the unregulated competitive equilibrium and the socially optimal

[15] Proving that any $x^e(\theta) > 0$ is socially excessive given X^e is immediate. However, the comparison between $x^e(\theta)$ and $x^*(\theta)$ for any θ is more tricky because (5) involves X^e while (9) involves $X^* < X^e$, and the differences between the second terms in each of these first-order conditions create potential ambiguity. In the numerical examples that we have explored, all banks use excessive short-term funding relative to the social optimum.

allocation under a parameterization whose details are easy to work out analytically. θ is assumed to be uniformly distributed over the $[0, 1]$ interval, the profit function is assumed to have the simple quadratic form $\pi(x, \theta) = (1 + \theta)x - \frac{1}{2}x^2$, and the product of the terms $\varepsilon(x, \theta)$ and $c(X)$ is jointly specified as $\varepsilon(x, \theta)c(X) = \alpha x X$, where α is a positive constant. Interior solutions can be guaranteed for $\alpha < 1/2$. The figure corresponds to the case with $\alpha = 0.15$. The horizontal axis reflects the values of θ and the vertical axis the unregulated equilibrium and the socially optimal allocations, $x^e(\theta)$ and $x^*(\theta)$, respectively. The unregulated equilibrium features systemic risk $X^e = 1.30$ and social welfare $W^e = 0.892$, while the social optimum features $X^* = 1.15$ and $W^* = 0.907$.

Summing up, systemic externalities associated with banks' short-term funding decisions create a positive wedge between the social and the private marginal costs of using short-term funding. Banks only internalize the implications of the funding choices for their own vulnerability to refinancing risk, disregarding their contribution to all other banks' systemic risk exposure and costs. Standard marginal reasoning when privately optimizing on short-term funding makes systemic risk in the unregulated equilibrium higher than socially optimal.

6. The Pigovian Tax: An Efficient Solution

As in the standard textbook discussion on the treatment of negative production externalities, the social efficiency of the competitive equilibrium can be restored by imposing a Pigovian tax, i.e., by taxing the activity causing the externality at a rate equal to the wedge between the social marginal cost and the private marginal cost of the activity (evaluated, if applicable, at the anticipated socially optimal allocation). In our case, this boils down to setting a flat tax per unit of short-term funding equal to

$$\tau^* = E_z(\varepsilon(x^*(z), z))c'(X^*). \tag{12}$$

Obviously, the introduction of a tax on short-term funding alters the first-order condition relevant for banks' optimization in the competitive equilibrium with taxes.

Formally, we can define a *competitive equilibrium with taxes* $\{\tau(\theta)\}$ as a pair $(\{x^{\tau}(\theta)\}, X^{\tau})$ satisfying

(i) $x^{\tau}(\theta) = \arg\max_x\{\pi(x,\theta) - \varepsilon(x,\theta)c(X^{\tau}) - \tau(\theta)x\}$ for all $\theta \in [0,1]$,

(ii) $X^{\tau} = \int_0^1 x^{\tau}(\theta)f(\theta)d\theta$.

In an interior equilibrium, the first-order conditions for the private optimality of each $x^{\tau}(\theta)$ imply

$$\pi_x(x^{\tau}(\theta),\theta) - \varepsilon_x(x^{\tau}(\theta),\theta)c(X^{\tau}) - \tau(\theta) = 0 \qquad (13)$$

for all $\theta \in [0,1]$. And it is immediate to see that the flat tax schedule $\tau(\theta) = \tau^*$, with the tax rate defined as in (12), makes (13) equivalent to (9), implying $(\{x^{\tau^*}(\theta)\}, X^{\tau^*}) = (\{x^*(\theta)\}, X^*)$ and, hence, implementing the socially optimal allocation as a competitive equilibrium.

To set the optimal tax rate τ^* properly, it is of course necessary that the regulator knows the functions that characterize the economy (including the density of the parameter θ that captures banks' heterogeneity) and is, hence, able to compute the socially optimal allocation that enters in (12). For the numerical example shown in figure 1, the flat Pigovian tax that implements the socially optimal allocation as a competitive equilibrium with taxes is $\tau^* = 0.17$.[16]

An important practical difficulty when regulating heterogeneous agents is that the particulars of the regulation applicable to each agent may depend on information that is private to the agent. This problem does not affect the efficient Pigovian tax τ^*, which is the same for all values of θ. The following proposition summarizes the key results of this section.

Proposition 2. *When banks differ in the marginal value they can extract from short-term funding, the socially optimal allocation can*

[16]Of course, this number depends on the underlying parameterization. Without touching other details of our simple example, which was not intended to provide realistic predictions, reducing the parameter α from 0.15 to, e.g., 0.015 reduces the efficient Pigovian tax to $\tau^* = 0.02$.

be reached as a competitive equilibrium by charging banks a flat Pigovian tax τ^* on each unit of short-term funding.

7. Quantity-Based and Ratio-Based Alternatives

Pigovian taxation is frequently described as a price-based solution to the regulation of externalities. Such a description emphasizes the capacity of the tax solution to decentralize the implementation of the desired allocation as a market equilibrium. The polar alternative is to go for a "centralized" quantity-based solution in which each regulated agent (bank) is directly mandated to choose its corresponding quantity (short-term funding) in the optimal allocation ($x^*(\theta)$ in the model).

In the context of our model, pure quantity-based regulation would require detailed knowledge by the regulator of individual marginal value of short-term funding for each bank (i.e., the derivatives $\pi_x(x, \theta)$ and $\varepsilon_x(x, \theta)$, which vary with θ and appear in (9)). Possibly due to the strong informational requirements that this implies, none of the alternatives for liquidity regulation considered in practice these days opt for directly setting individualized quantity prescriptions such as $x^*(\theta)$.

The alternatives to Pigovian taxes actually under discussion are ratio-based regulations, i.e., regulations that consist of forcing banks to have some critical accounting ratios above or below some regulatory minima or maxima. To be sure, some proposals include making the regulatory limits functions of individual characteristics of each bank—such as size, interconnectedness, lack of substitutability, global (cross-jurisdictional) activity, and complexity—but none of the considered characteristics seem targeted to control for the heterogeneity in banks' capacity to extract value from short-term funding.[17] These qualifiers can be rather rationalized as an attempt to capture what, in an extension discussed in section 9, we describe as the *systemic importance* of each bank (the relative importance of the contribution of its short-term funding to the systemic risk measure X).

[17]See, for example, the press release on "Measures for Global Systemically Important Banks Agreed by the Group of Governors and Heads of Supervision" issued by the BCBS on June 25, 2011 (see www.bis.org/press/p110625.htm).

The most seriously considered ratio-based alternatives for the regulation of liquidity are those put forward by the new Basel III agreement (see Basel Committee on Banking Supervision 2010). The agreement introduces two new regulatory ratios: a *liquidity coverage ratio*, similar in format and spirit to one already introduced by the Financial Services Authority in the United Kingdom in October 2009, and a more innovative *net stable funding ratio*. To facilitate the discussion, we analyze each of these instruments as if it were introduced in isolation, starting with the last one, whose potential effectiveness for the regulation of funding maturity is somewhat less ambiguous.

7.1 A Stable Funding Requirement

The net stable funding requirement calls banks to hold some accounting ratio of "stable funding" (i.e., equity, customer deposits, and other long-term or "stable" sources of funding) to "non-liquid assets" above some regulatory minimum. To translate this to our model, where banks' asset composition and stable sources of funding have been so far treated as exogenously fixed, we can think of this requirement as equivalent to imposing an upper limit $\overline{x}$ to the short-term debt that each bank can issue. In a more general version of our model, the effective upper limit applicable to each bank could be considered affected by prior decisions regarding the maturity and liquidity structure of the bank's assets, its retail deposits base, its level of capitalization, etc. But here, for simplicity, one can think of changes in these decisions as possible interpretations for the comparative statics of $\overline{x}$.

The introduction of a minimum stable funding requirement has then the implication of adding an inequality constraint of the type $x \leq \overline{x}$ to the private optimization problem of the banks. Formally, a *competitive equilibrium with a stable funding requirement* parameterized by $\overline{x}$ can be defined as a pair $(\{x^{\overline{x}}(\theta)\}, X^{\overline{x}})$ satisfying

(i) $x^{\overline{x}}(\theta) = \arg\max_{x \leq \overline{x}}\{\pi(x,\theta) - \varepsilon(x,\theta)c(X^{\overline{x}})\}$ for all $\theta \in [0,1]$,
(ii) $X^{\overline{x}} = \int_0^1 x^{\overline{x}}(\theta) f(\theta) d\theta$.

Since the preference for short-term funding is strictly increasing in θ, we may have up to three possible configurations of equilibrium.

For $\overline{x} \geq x^e(1)$, the stable funding requirement will not be binding for any bank (since $\theta = 1$ identifies the banks with the highest incentives to use short-term funding), and the equilibrium will then coincide with the unregulated competitive equilibrium characterized in section 4. For $\overline{x} \leq x^e(0)$, the stable funding requirement will be binding for all banks (since $\theta = 1$ identifies the banks with the lowest incentives to use short-term funding), implying $x^{\overline{x}}(\theta) = \overline{x} < x^e(\theta)$ for all θ and, hence, $X^{\overline{x}} = \overline{x} < X^e$. Finally, for $\overline{x} \in (x^e(0), x^e(1))$, the stable funding requirement will be binding for at least the banks with the largest θs and perhaps for all banks. To see the latter, notice that inducing the limit choice of $x^{\overline{x}}(\theta) = \overline{x} < x^e(\theta)$ to the banks with relatively large θs will push $X^{\overline{x}}$ below X^e, but this, in turn, will push the banks with relatively low θs into choices of $x^{\overline{x}}(\theta) > x^e(\theta)$, possibly (but not necessarily) inducing some or even all of them to also hit the regulatory limit $\overline{x}$.

It is then obvious that, in general, a sufficiently tight stable funding requirement $\overline{x} < x^e(1)$ can reduce the equilibrium measure of aggregate systemic risk $X^{\overline{x}}$ relative to the unregulated equilibrium X^e, thus moving it closer to its value in the socially optimal allocation X^*. The induced allocation will, however, be necessarily inefficient. The reason for this is that the reduction in the activities that generate negative externalities comes at the cost of distorting the allocation of short-term funding across bank classes: (i) constraining the banks with relatively higher valuation for short-term funding to the common upper limit $\overline{x}$, and (ii) encouraging the banks with relatively low valuation for short-term funding to use more of it than would be socially optimal (since they will choose $x^{\overline{x}}(\theta) > x^e(\theta) > x^*(\theta)$). In fact, there is no guarantee that introducing a $\overline{x}$ that simply brings $X^{\overline{x}}$ closer to X^* improves, in welfare terms, over the unregulated equilibrium.

PROPOSITION 3. *A binding net stable funding requirement will affect the measure of aggregate systemic risk X in the same direction as the efficient arrangement (i.e., will reduce X), but it will also redistribute short-term funding inefficiently from banks that value it more to banks that value it less, so that the socially optimal allocation cannot be reached and the improvement in social welfare is not guaranteed.*

Assuming a solution in which $\overline{x}$ is binding for some but not all banks, the socially optimal choice of $\overline{x}$ can be defined as follows:

$$\overline{x}^{SB} = \arg \max_{(\overline{x}, X^{\overline{x}})} \int_0^{\overline{\theta}} [\pi(y(\theta, X^{\overline{x}}), \theta) - \varepsilon(y(\theta, X^{\overline{x}}), \theta) c(X^{\overline{x}})] f(\theta) d\theta$$
$$+ \int_{\overline{\theta}}^1 [\pi(\overline{x}, \theta) - \varepsilon(\overline{x}, \theta) c(X^{\overline{x}})] f(\theta) d\theta$$
$$\text{s.t.:} \quad \int_0^{\overline{\theta}} y(\theta, X^{\overline{x}}) f(\theta) d\theta + \overline{x}[1 - F(\overline{\theta})] = X^{\overline{x}}, \tag{14}$$

where the function $y(\theta, X)$ is defined as in (4) and $F(\theta)$ is the cumulative distribution function associated with $f(\theta)$. This problem identifies the "second-best" allocation attainable if the only available instrument for liquidity regulation is the stable funding requirement $\overline{x}$, which is assumed to be binding only for the banks with $\theta \in (\overline{\theta}, 1]$, where $\overline{\theta}$ is implicitly defined by $y(\overline{\theta}, X^{\overline{x}}) = \overline{x}$.

To solve the problem in (14), the simplest approach is to treat the constraint as a definition of the aggregate systemic risk $X^{\overline{x}}$ induced by the choice of $\overline{x}$, so that $\overline{x}$ is effectively the only decision variable. The first-order condition for an optimal interior solution in this variable will then require that the full differential of the objective function above with respect to $\overline{x}$ is made equal to zero at the optimum $\overline{x}^{SB}$. Such full differential of the objective function will generally have terms related to (i) the direct effects of a marginal variation in $\overline{x}$ on the integrands, (ii) the effects of a marginal variation in $\overline{x}$ on $\overline{\theta}$ (and through it on social welfare), and (iii) the effects of a marginal variation in $\overline{x}$ on $X^{\overline{x}}$ (and through it on social welfare). However, it turns out that the effects of type (i) are only relevant in the second term of the objective function (i.e., for the net value generated by banks for which the stable funding requirement is binding), (ii) is zero, and the effects of type (iii) channelled via $y(\theta, X^{\overline{x}})$ for the banks for which the stable funding requirement is not binding are also zero by virtue of (4).[18]

[18]This last effect is simply the particularization of the envelope theorem to the individual optimization decisions of unconstrained banks.

Eventually, the only non-zero terms resulting from the full differentiation of social welfare with respect to $\overline{x}$ give rise to the following first-order condition for an interior solution to the maximization problem in (14):

$$\int_{\overline{\theta}}^{1} [\pi_x(\overline{x},\theta) - \varepsilon_x(\overline{x},\theta)c(X^{\overline{x}})]f(\theta)d\theta - E_\theta(\varepsilon(x^{\overline{x}}(\theta),\theta))c'(X^{\overline{x}})\frac{dX^{\overline{x}}}{d\overline{x}} = 0, \tag{15}$$

where

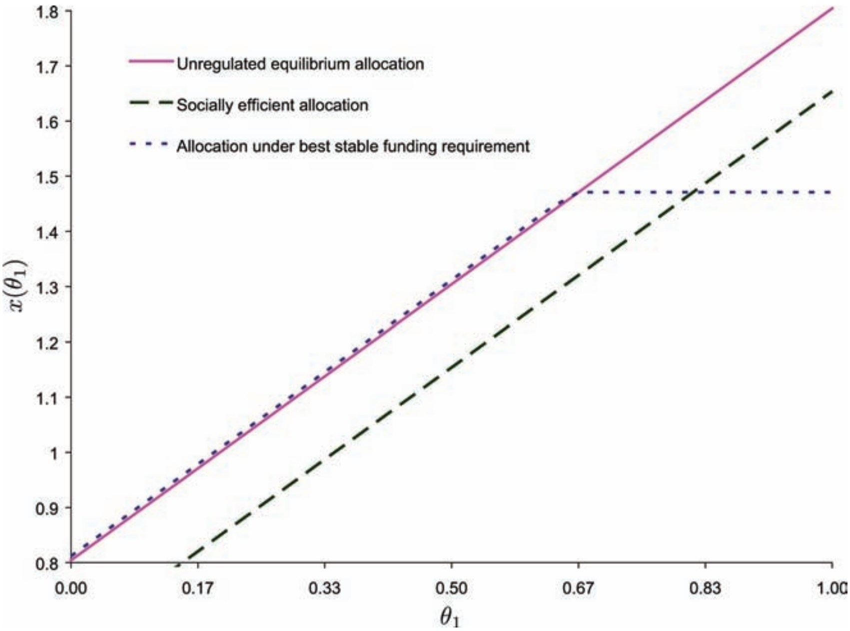
$$\frac{dX^{\overline{x}}}{d\overline{x}} = \frac{1 - F(\overline{\theta})}{1 - \int_0^{\overline{\theta}} y_X(\theta, X^{\overline{x}})f(\theta)d\theta} \in [0,1]. \tag{16}$$

The first term on the left-hand side of (15) collects the direct effects of a marginal change in $\overline{x}$ on the expected NPV generated by the banks for which the stable funding requirement is binding. This term is positive because, for given $X^{\overline{x}}$, relaxing the constraint associated with $\overline{x}$ would produce more NPV at each constrained bank. The second term captures the impact of the requirement $\overline{x}$ on aggregate systemic risk $X^{\overline{x}}$ and, through it, on the negative externalities suffered by all banks. This term is clearly negative and the only reason why introducing a binding $\overline{x}^{SB}$ may be possibly optimal in a second-best sense. As reflected in (16), $dX^{\overline{x}}/d\overline{x}$ will typically be lower than one because moving $\overline{x}$ will directly reduce the systemic risk generated by the constrained banks only, while indirectly encouraging (due to the very reduction in $X^{\overline{x}}$) the use of short-term funding by the unconstrained banks.[19]

To gain further intuition on the second-best nature of the tradeoffs behind the choice of $\overline{x}^{SB}$, it is convenient to compare (15) with the condition for first-best efficiency in (9). First, (9) applies pointwise, defining an efficient $x^*(\theta)$ for each θ; in contrast, (15) determines a common upper limit $\overline{x}$ to the short-term funding decisions of all banks. Hence, in (15) the costs and benefits of marginally moving $\overline{x}$ are "averaged" over all the θs. The terms in the integral that appears in (15) resemble the first two terms on the left-hand side of

[19] The fact that $y(\theta, X^{\overline{x}})$ is decreasing in $X^{\overline{x}}$ follows directly from applying the implicit function theorem on (4).

Figure 2. Equilibrium Allocation under the Best Stable Funding Requirement



(9), but the ones "averaged" here correspond to the set of high θs only, for which the requirement $\overline{x}$ is binding—the remaining banks are not directly affected by $\overline{x}$. The second term in (15) and the third in (9) reflect the marginal externality caused by $\overline{x}$ and each $x^*(\theta)$, respectively. The relevant difference regarding these terms is due to the presence of $dX^{\overline{x}}/d\overline{x}$ (typically lower than one) in (15).

Intuitively, the stable funding requirement emerges as just too rough a way to deal with the underlying systemic externalities. It hinges on constraining in a fully indiscriminate manner the decisions made by the banks that, under the source of heterogeneity considered here, are precisely those that can generate more value from short-term funding. And it has the undesirable (though possibly quantitatively small) side effect of encouraging those that generate less value from short-term funding to use more of it. In the numerical example depicted in figure 2, built on the same assumptions as figure 1, we have that $\overline{x}^{SB} = 1.47$, which means the best stable funding requirement is binding for the banks in the top 34 percent quantile of the distribution of θs. For lower values of θ, the resulting allocation involves some $x^{\overline{x}}(\theta)$ slightly above those in the unregulated equilibrium, $x^e(\theta)$, and the whole second-best allocation attainable with $\overline{x}^{SB}$ is quite distant from the socially efficient one, $x^*(\theta)$. The best stable funding requirement leads to a social welfare of $W^{\overline{x}} = 0.896$ (only 0.4 percent larger than in the unregulated equilibrium), while the optimal flat Pigovian tax reaches the first-best level $W^* = 0.907$ (1.6 percent more than in the unregulated equilibrium). The best stable funding requirement reduces aggregate systemic risk by 3.9 percent, while the optimal Pigovian tax reduces it by 11.5 percent.

7.2 A Liquidity Requirement

The liquidity coverage ratio introduced by Basel III requires banks to back their use of short-term funding with the holding of high-quality liquid assets, i.e., assets that could be easily sold, presumably at no fire-sale loss, in case of a crisis. In its original description, this requirement responds to the motivation of providing each bank with its own liquidity buffer, which, presumably, might also expand the liquidity available in the system in case of a crisis (on top of that possibly provided by the lender of last resort).

Specifically, it is proposed that banks estimate the refinancing needs that they would accumulate if the functioning of money markets or other conventional borrowing sources were disrupted for some specified period (one month) and keep enough high-quality liquid assets so as to be able to confront the situation with their sale.[20] Qualifying assets would essentially be cash, central bank reserves, and treasury bonds.

How can we capture this requirement in the context of our model? Leaving details aside, the liquidity requirement can be seen as a requirement to back some minimal fraction $\phi \leq 1$ of each bank's short-term funding x with the holding of qualifying liquid assets m, thereby introducing the constraint $m \geq \phi x$. Additionally, the impact

[20]Another option would be to post them as collateral at the central bank's discount window.

of m on the bank's objective function could be taken into account by considering the following extended value function:

$$v(x, m, X, \theta) = \pi(x - m, \theta) - \varepsilon(x - m, \theta)c(\widehat{X}) - \delta m, \qquad (17)$$

where

$$\widehat{X} = \int_0^1 [x(\theta) - m(\theta)]f(\theta)d\theta, \qquad (18)$$

and $\delta = r_b - r_m \geq 0$ is the difference between the bank's short-term borrowing rate r_b and the yield r_m of the qualifying liquid assets. This formulation credits for both the individual and the systemic "buffering" role of the liquid assets by making each bank's individual vulnerability factor $\varepsilon(x - m, \theta)$ a function of its *net* short-term funding and by redefining the systemic risk measure $\widehat{X}$ as the result of aggregating banks' net short-term funding positions.

The other terms in (17) capture the NPV generated in the absence of a systemic crisis. Our formulation is based on assuming that the former function $\pi(x, \theta)$ only captured the NPV generated by the bank's core lending or investment activity, which does not include investing in the qualifying liquid assets. The new first argument of $\pi(x - m, \theta)$ is justified by the fact that if a part m of the resources obtained as short-term funding x is invested in liquid assets, the net amount available for core banking activities becomes $x - m$. We assume that the funds m invested in liquid assets yield a risk-free rate r_m but have a cost equal to the bank's short-term borrowing rate $r_b \geq r_m$, in which, for simplicity, we will treat r_b as exogenous (and equal for all banks).[21] Hence the spread $\delta = r_b - r_m \geq 0$ acts as a direct net cost of holding liquid assets.

[21] In principle, r_b should depend on each bank's risk of insolvency as assessed by market participants. In the logic of the model, one might think of the spread as directly connected to $\varepsilon(x - m, \theta)$ and perhaps also to $c(X)$. However, the exogeneity of r_b can be justified by alluding either to informational asymmetries or safety-net guarantees that effectively make investors not discriminate across banks. Additionally, one could also argue that the main driver of banks' insolvency risk is asset risk, from which we are abstracting.

In this extended framework, social welfare can be written as

$$W(\{x(\theta), m(\theta)\}) = \int_0^1 [\pi(x(\theta) - m(\theta), \theta) - \varepsilon(x(\theta) - m(\theta), \theta)c(\widehat{X}) - \delta m(\theta)]f(\theta)d\theta, \tag{19}$$

where the presence of $-\delta m(\theta)$ implies considering banks' direct costs of holding liquidity as a deadweight loss. This formulation is consistent with having assumed that investors provide (short-term) funding to the banks at competitive market rates and thus make zero NPV when doing so. In this context, $\delta > 0$ can be interpreted as the premium that compensates investors for (unmodeled) utility losses derived from either the risk or the lower liquidity of their investment in bank liabilities.

A *competitive equilibrium with a liquidity requirement* parameterized by ϕ can be defined as a pair $(\{(x^\phi(\theta), m^\phi(\theta))\}, \widehat{X}^\phi)$ satisfying

(i) $(x^\phi(\theta), m^\phi(\theta)) = \arg\max_{m \geq \phi x}\{\pi(x-m, \theta) - \varepsilon(x-m, \theta)c(\widehat{X}^\phi) - \delta m\}$ for all $\theta \in [0, 1]$,
(ii) $\widehat{X}^\phi = \int_0^1 (x^\phi(\theta) - m^\phi(\theta))f(\theta)d\theta$.

Before continuing, it may be worth commenting on how equilibrium looks in the special limit case with $\phi = 1$. Obliging banks to match each unit of short-term funding with one unit of liquid assets implies that banks are left with no funds to expand their value-creating activities. In other words, banks will have $\widehat{x} = x - m = 0$ for any possible choice of x. However, any $x > 0$ will imply $m = x > 0$, and hence paying the corresponding direct cost $\delta m > 0$ of holding liquidity, so banks will optimally fix $x = m = 0$. Of course, aggregate systemic risk will be zero, i.e., $\widehat{X}^{\phi=1} = 0$, but the welfare generated by the banks, as measured by (19), will also be zero. In what follows we will focus on the case with $\phi < 1$.

In general the liquidity requirement can be taken as binding (necessarily so if $\delta > 0$ and binding without loss of generality if $\delta = 0$). For $\phi < 1$ it is possible and convenient to reformulate banks' optimization problem in terms of their choice of *net* short-term funding $\widehat{x}(\theta) = x(\theta) - m(\theta)$ only, since the binding liquidity constraint allows us to write $m(\theta)$ as $\frac{\phi}{1-\phi}\widehat{x}(\theta)$ and eliminate it as an independent decision variable. Hence, equilibrium can be redefined as a pair $(\{\widehat{x}^{\phi}(\theta)\}, \widehat{X}^{\phi})$ satisfying

(i) $\widehat{x}^{\phi}(\theta) = \arg\max_{\widehat{x}}\{\pi(\widehat{x},\theta) - \varepsilon(\widehat{x},\theta)c(\widehat{X}^{\phi}) - \frac{\delta\phi}{1-\phi}\widehat{x}\}$ for all $\theta \in [0,1]$,
(ii) $\widehat{X}^{\phi} = \int_0^1 \widehat{x}^{\phi}(\theta)f(\theta)d\theta$.

We will proceed with the analysis by looking first at the case in which the net cost of holding liquid assets is zero ($\delta = 0$) and then at the case in which it is positive ($\delta > 0$).

7.2.1 The Case in which Holding Liquidity Is Costless ($\delta = 0$)

The following proposition establishes a somewhat shocking result for the relevant case in which the spread δ is zero (roughly the case in "normal times," when banks are perceived as essentially risk-free borrowers).

PROPOSITION 4. *With $\delta = 0$, the competitive equilibrium with a liquidity requirement $\phi < 1$ involves the same amount of net short-term funding and, hence, the same level of systemic risk as the unregulated equilibrium. That is, it involves $x^{\phi}(\theta) - m^{\phi}(\theta) = x^{e}(\theta)$ and $\widehat{X}^{\phi} = X^{e}$.*

The proof of this proposition follows immediately from the equivalence, when $\delta = 0$, between the equilibrium conditions for $(\{\widehat{x}^{\phi}(\theta)\}, \widehat{X}^{\phi})$ and those for $(\{x^{e}(\theta)\}, X^{e})$ (see section 4). Hence, the only effect of the liquidity requirement relative to the unregulated equilibrium is to induce an artificial demand $M^{\phi} = \frac{\phi}{1-\phi}E_{\theta}(x^{e}(\theta))$ for the qualifying liquid assets and a spurious increase in banks' *gross* short-term funding, which becomes $E_{\theta}(x^{\phi}(\theta)) = E_{\theta}(x^{e}(\theta)) + M = \frac{1}{1-\phi}E_{\theta}(x^{e}(\theta))$.

Therefore, when the direct net cost δ of each unit of liquidity that the requirement forces banks to hold is zero (not implausible in "normal times"), the liquidity coverage ratio totally fails to bring the equilibrium allocation any closer to the social optimum than in the unregulated scenario. Banks respond to regulation by increasing their short-term funding and their liquidity holding so as to make their net short-term funding as high as in the unregulated equilibrium. The artificial demand for high-quality liquid assets may imply that liquid assets, kept somewhere else in the financial system (e.g., money market mutual funds) prior to imposing the ratio, end up kept by banks after imposing the ratio. However, the (net) systemic risk generated by the banks will not change.

7.2.2 The Case in which Holding Liquidity Is Costly ($\delta > 0$)

When the direct net unit cost of holding liquidity, δ, is positive, the implications are quite different. The conditions that characterize $(\{\widehat{x}^{\phi}(\theta)\}, \widehat{X}^{\phi})$ become analogous to those that characterize a competitive equilibrium with taxes in which $\tau(\theta) = \frac{\delta\phi}{1-\phi}$ (see section 6). This allows us to directly state the following proposition.

PROPOSITION 5. *With $\delta > 0$, the competitive equilibrium with a liquidity requirement $\phi < 1$ involves the same individual net short-term funding decisions and aggregate systemic risk as a competitive equilibrium with a tax on short-term funding with rate $\tau(\theta) = \frac{\delta\phi}{1-\phi}$ for all θ.*

For a given $\delta > 0$, the implicit tax rate $\delta\phi/(1-\phi)$ described above moves from zero to infinity as the liquidity requirement ϕ moves from zero to one. Thus the regulator can *seemingly* replicate the effects of any flat tax τ (including the efficient Pigovian tax τ^* of section 6) by setting $\phi = \frac{\tau}{\delta+\tau}$. However, banks' demand for the qualifying liquid assets would be $m^{\phi}(\theta) = \frac{\phi}{1-\phi}\widehat{x}^{\phi}(\theta) = \frac{\tau}{\delta}x^{\tau}(\theta)$ (implying an aggregate demand $M^{\phi} = \frac{\tau}{\delta}X^{\tau}$) and their *gross* short-term funding would be $x^{\phi}(\theta) = x^{\tau}(\theta) + m^{\phi}(\theta) = \frac{\delta+\tau}{\delta}x^{\tau}(\theta) > x^{\tau}(\theta)$ (implying $X^{\phi} = E_{\theta}(x^{\phi}(\theta)) = X^{\tau} + M^{\phi} = \frac{\delta+\tau}{\delta}X^{\tau} > X^{\tau}$ at the aggregate level). Importantly, the total direct net costs of holding liquidity would cause a deadweight loss of $\delta m^{\phi}(\theta) = \tau x^{\tau}(\theta)$ to each bank. Not surprisingly, the aggregate deadweight loss $\delta M^{\phi} = \tau X^{\tau}$ equals the tax revenue that the replicated tax on short-term funding could have raised.

The presence of the deadweight loss τX^{τ} implies that the liquidity requirement that *seemingly* replicates the Pigovian solution $\left(\phi^* = \frac{\tau^*}{\delta+\tau^*}\right)$ is not socially efficient.

PROPOSITION 6. *With $\delta > 0$, replicating the net short-term funding allocation and aggregate systemic risk of the efficient allocation using a liquidity requirement $\phi^* = \frac{\tau^*}{\delta+\tau^*}$ is feasible but entails a deadweight loss $\tau^* X^* > 0$.*

Actually, ϕ^* will not generally be optimal even from a second-best perspective, except in the non-generic situation in which the efficient Pigovian tax τ^* happens to be at a critical point of the Laffer curve τX^τ. This is because moving the liquidity requirement marginally away from ϕ^* (in one direction) will reduce the deadweight loss δM^ϕ, while other components of social welfare will not change (since they are maximized precisely with $\phi = \phi^*$).

For a given spread $\delta > 0$, the socially optimal liquidity requirement will be some $\phi^{SB} = \frac{\tau^{SB}}{\delta+\tau^{SB}}$ whose associated implicit tax rate τ^{SB} satisfies

$$\tau^{SB} = \arg\max_{\tau\geq 0} \int_0^1 [\pi(x^\tau(\theta),\theta) - \varepsilon(x^\tau(\theta),\theta)c(X^\tau) - \tau x^\tau(\theta)] f(\theta) d\theta$$

$$\text{s.t.:} \quad x^\tau(\theta) = \arg\max_x \pi(x,\theta) - \varepsilon(x,\theta)c(X^\tau) - \tau x \text{ for all } \theta$$

$$\int_0^1 x^\tau(\theta) f(\theta) d\theta = X^\tau. \tag{20}$$

This formulation of the optimization problem exploits the analogy explained above, which conveniently allows us to write the deadweight loss suffered by each bank as $\tau x^\tau(\theta)$, which is actually independent of δ and will end up making the solution in terms of τ^{SB} also independent of δ. The constraints in the optimization problem are simply the conditions that define an equilibrium with a tax τ on short-term funding (see section 6).

Typically, the optimal liquidity requirement ϕ^{SB} will be inferior to ϕ^*, implying more aggregate systemic risk than in the first-best allocation. The intuition for this is that moving away from the unregulated equilibrium allocation by increasing ϕ will typically monotonically increase the aggregate deadweight loss δM^ϕ, while the remaining marginal benefits of moving towards the first-best allocation decline towards zero as ϕ approaches ϕ^*. In fact, for the parameterization behind figures 1 and 2, the optimal liquidity requirement ϕ^{SB} is just zero, meaning that the deadweight losses associated with forcing banks to hold liquidity do not compensate, for any level of ϕ, the gains from the reduction in systemic risk.[22]

Interestingly, the writing of the problem as in (20) makes clear that τ^{SB} does not depend on δ, implying that the total variation of $\phi^{SB} = \frac{\tau^{SB}}{\delta+\tau^{SB}}$ with respect to δ is just given by the partial derivative

$$\frac{\partial \phi^{SB}}{\partial \delta} = \frac{-\tau^{SB}}{(\delta + \tau^{SB})^2} < 0. \tag{21}$$

Hence, if the regulator wants to implement an interior second-best allocation as defined above (or to seemingly replicate the efficient Pigovian tax), it should be ready to move the imposed liquidity requirement ϕ^{SB} (or ϕ^*) in response to the fluctuations in the spread δ. In practice, moving ϕ and the implied adjustments in quantities may be a source of trouble. On the one hand, authorities will have to be effective in changing ϕ in due course. On the other hand, frequent and sudden changes in ϕ might produce volatility in the demand for the liquid assets included in M^ϕ. This might be especially so if δ approaches zero, in which case the prescriptions for ϕ^{SB} (or ϕ^*) imply that M^ϕ would tend to infinity.

With potentially large variations in the demand for liquid assets, the reference to possible general equilibrium implications is inexcusable. Our results are definitely valid if the supply of liquid assets is sufficiently elastic at the rate r_m used in the definition of the spread δ. However, if changes in liquidity requirements have an impact on the equilibrium value of r_m due to a more general interaction between the demand and supply for liquid assets, then the required analysis may well exceed the scope of this paper. This is especially so if r_m is also the reference risk-free rate in the economy, because then the net present values captured by our reduced-form value functions might be expected to change with r_m and the whole analysis in (at least) this section should be redone using a somewhat more structural formulation.

---

[22]In our example, the liquidity requirement that replicates the effects of the optimal Pigovian tax ($\tau^* = 0.17$) has deadweight costs equivalent to 22 percent of the value generated by banks in the unregulated equilibrium. These costs are more than ten times the increase in welfare that the pure Pigovian solution might produce relative to the unregulated equilibrium.

8. Risk Shifting and the Case for Quantity Regulation

In this section we extend the model to address formally one of the main criticisms to the proposal of a Pigovian approach to liquidity risk regulation. Such criticism is based on the "robustness" of the price-based approach to modeling mistakes and, specifically, to the possibility of having some "crazy" or just particularly risk-inclined banks that, for the sake of expanding their risky lending, are willing to pay large amounts of the established tax so as to use large amounts of short-term funding.

In our baseline formulation, banks that like to take more short-term funding are those that can extract more expected NPV from it. In such a formulation, the considered dimension of heterogeneity makes banks with larger θ essentially more valuable—privately and, if properly regulated, also socially. We will now denote that dimension of heterogeneity by θ_1 and introduce a second dimension of heterogeneity, $\theta_2 \in [0, 1]$, intended to capture differences in banks' inclination towards risk taking.[23] The joint distribution of (θ_1, θ_2) will be described by the density function $f(\theta_1, \theta_2)$.

To capture heterogeneity in banks' risk-shifting inclinations formally, we are going to treat θ_2 as a parameter that determines the fraction of the losses incurred by a bank during a crisis which are *not* internalized by its owners but passed (without compensation) to other stakeholders (e.g., the deposit insurer). We then assume each bank, when privately deciding on x, only considers the fraction $1-\theta_2$ of $\varepsilon(x, \theta)c(X)$ as an expected value loss, leaving the remaining fraction θ_2 to other stakeholders. Hence, the social welfare measure $W(\{x(\theta)\})$ must now explicitly consider, in addition to the NPV

[23] The literature has identified several sources of such differences. Corporate governance arrangements may affect the severity of the conflicts of interest between shareholders and debtholders, making the former more or less capable to ex post expropriate the former by shifting risk (Jensen and Meckling 1976). In the case of banks, risk-shifting problems are exacerbated by the existence of safety-net guarantees (e.g., deposit insurance) provided at risk-insensitive rates. In such a setup, banks' charter values reduce excessive risk taking (Keeley 1990). Capital requirements (especially if risk based) generally improve the alignment of incentives between the bankers and other stakeholders (Holmstrom and Tirole 1997) and can specifically attenuate the risk-shifting problem (Hellmann, Murdock, and Stiglitz 2000; Repullo 2004).

appropriated by the bank owners, the losses $-\theta_2\varepsilon(x,\theta)c(X)$ passed on to other bank stakeholders.

So the new objective function for banks is

$$v(x, X, \theta_1, \theta_2) = \pi(x, \theta_1) - (1 - \theta_2)\varepsilon(x, \theta_1)c(X), \tag{22}$$

while social welfare is given by

$$W(\{x(\theta_1, \theta_2)\}) = \int_0^1 \int_0^1 [v(x(\theta_1, \theta_2), X, \theta_1, \theta_2) \\ - \theta_2\varepsilon(x(\theta_1, \theta_2), \theta_1)c(X)]f(\theta_1, \theta_2)d\theta_1 d\theta_2, \tag{23}$$

where

$$X = g(\{x(\theta_1, \theta_2)\}) = \int_0^1 \int_0^1 x(\theta_1, \theta_2)f(\theta_1, \theta_2)d\theta_1 d\theta_2. \tag{24}$$

Plugging (22) into (23), social welfare can be written as

$$W(\{x(\theta_1, \theta_2)\}) = \int_0^1 \int_0^1 [\pi(x(\theta_1, \theta_2), \theta_1) - \varepsilon(x(\theta_1, \theta_2), \theta_1)c(X)] \\ \times f(\theta_1, \theta_2)d\theta_1 d\theta_2, \tag{25}$$

which is conceptually identical to (3).

8.1 Gambling as the Sole Source of Heterogeneity

To highlight our key argument, suppose that the variation due to θ_1, whose implications we have already discussed in prior sections, is shut down by fixing $\theta_1 = \overline{\theta}_1$ for all banks. So residual bank heterogeneity is due to θ_2 only. How is the unregulated equilibrium determined? And the socially optimal allocation? How do they differ? How should $x(\theta_2)$ be regulated?

Without restating all the relevant definitions (which will follow mechanically from the adaptation of those already presented for the baseline model), the answers to these questions can be found by comparing the first-order conditions satisfied by bank decisions, $x^{ee}(\theta_2)$, and the systemic risk measure, X^{ee}, in an interior unregulated equilibrium, with the conditions satisfied by their counterparts, $x^{**}(\theta_2)$

and X^{**}, in an interior socially optimal allocation. Similarly to (5), the unregulated equilibrium objects satisfy

$$\pi_x(x^{ee}(\theta_2), \overline{\theta}_1) - (1 - \theta_2)\varepsilon_x(x^{ee}(\theta_2), \overline{\theta}_1)c(X^{ee}) = 0, \qquad (26)$$

while, similarly to (9), in the socially optimal allocation we must have

$$\pi_x(x^{**}(\theta_2), \overline{\theta}_1) - \varepsilon_x(x^{**}(\theta_2), \overline{\theta}_1)c(X^{**}) - E_z(\varepsilon(x^{**}(z), \overline{\theta}_1))c'(X^{**}) = 0, \qquad (27)$$

in both cases for all θ_2. From these conditions, it is immediate to conclude that $x^{ee}(\theta_2)$ is increasing in θ_2 (that is, banks with greater risk-shifting inclinations tend to use more short-term funding) while $x^{**}(\theta_2)$ is independent of θ_2 and, hence, equal to a constant $\overline{x}^{**}$ (since, for any given x, θ_2 determines the distribution of value across bank stakeholders but not the total marginal value of short-term funding).

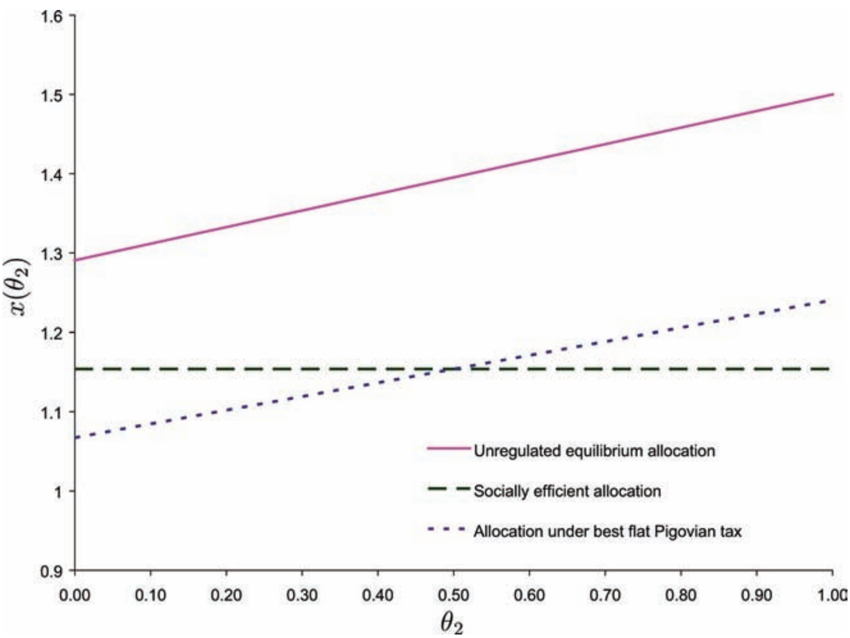
By simple comparison of the two sets of conditions, it is now obvious that the efficient Pigovian tax schedule is

$$\tau^{**}(\theta_2) = \theta_2\varepsilon_x(x^{**}(\theta_2), \overline{\theta}_1)c(X^{**}) + E_z(\varepsilon(x^{**}(z), \overline{\theta}_1))c'(X^{**}), \quad (28)$$

where the first term is new relative to (12) and reflects that risk-shifting incentives produce additional discrepancies between the private and social costs of expanding banks' short-term funding. In contrast to the pure systemic externality term (identical to what we had in the baseline model), the first term depends on θ_2. Hence, the efficient Pigovian tax schedule is not flat and cannot be enforced without detailed knowledge of each bank's risk-shifting inclination. A flat tax on short-term funding will not implement the first-best allocation.

Now, however, proper quantity regulation can do a great job. Specifically, a net stable funding requirement that effectively imposes the first-best quantity $\overline{x}^{**}$ as a limit to each bank's use of short-term funding would implement the first best. It is easy to see that the regulatory constraint will be binding for all θ_2. As for liquidity requirements, the rather negative conclusions obtained in the baseline analysis would still apply: with $\delta = 0$, a liquidity requirement

Figure 3. Allocations when Banks Differ in Gambling Incentives



is as ineffective as it was there, while with $\delta > 0$ its effect is very similar to (but with worse welfare properties than) a flat tax on short-term funding. And a flat tax on short-term funding is not a good solution in this environment!

Our conclusions can then be summarized as follows:

PROPOSITION 7. *If gambling incentives constitute the only source of heterogeneity across banks, a stable funding requirement* $\overline{x} = \overline{x}^{**}$ *implements the socially efficient allocation, while no flat-rate tax on short-term funding can do it. A liquidity requirement has the same shortcomings as in the baseline model and is, then, either ineffective (if* $\delta = 0$*) or very similar (but with larger deadweight costs) to the flat-rate tax solution (if* $\delta > 0$*).*

As a graphical illustration of these results, figure 3 develops an example based on extending and modifying the parameterization used in previous figures. We fix $\theta_1 = 1$ and introduce θ_2 as in (22) under the assumption that it is uniformly distributed over the interval $[0, 1]$. The new figure depicts, as a function of θ_2, the unregulated equilibrium allocation $x^{ee}(\theta_2)$, the socially optimal allocation $x^{**}(\theta_2)$ (which can be attained by limiting banks' short-term funding to $\overline{x}^{**} = 1.154$), and the allocation that would arise under the best flat tax on short-term funding.[24] In this example, social welfare in the unregulated equilibrium is 95.4 percent of its first-best level, while under the flat Pigovian tax it reaches 99.8 percent of the first-best level. This good second-best performance is due to the fact that, as illustrated in the figure, the downward shift in the use of short-term funding by all banks induced by the flat tax (which corrects for the "average" externalities of the unregulated equilibrium) is of relatively larger importance than the "differences" in externalities produced by the heterogeneity in θ_2.

8.2 Generalizing the Analysis

The analysis of the general case in which both θ_1 and θ_2 exhibit significant variation across banks is complicated and unlikely to yield clear-cut analytical results, if anything because first-best efficiency will not be generally attainable using instruments that are not explicitly contingent on θ_1 or θ_2. The analysis of simple instruments will necessarily be based on their second-best performance, which will have to be checked numerically for the general case.

Using a continuity argument and building on the polar cases already analyzed above, we can say that a flat tax on short-term funding will tend to perform better than a stable funding requirement if θ_1 is the dominant source of heterogeneity, i.e., if it has ample variation and, specifically, sufficient density at its upper tail, producing sufficiently many banks with value-generating motives to use short-term funding at a larger scale. The opposite will be true if θ_2 is the dominant source of variation, in this case producing sufficiently many banks whose main reason for wanting to use short-term funding at large scale is risk shifting. For instance, if the banking system had a small group of gambling banks and an ample majority

[24] In our example, the trade-offs involved in determining the best flat tax rate $\overline{\tau}^{SB}$ are linear, so $\overline{\tau}^{SB}$ (which equals 0.26) coincides with the average value of the efficient Pigovian schedule $\tau^{**}(\theta_2)$ characterized in equation (28).

of non-gambling banks, a stable funding requirement might be helpful to control the otherwise excessive short-term funding that the former would like to use.

But continuing with these easy-to-visualize examples, one can also anticipate possible advantages from combining the instruments. Suppose, in particular, that there is a group of banks which are essentially diverse in θ_1 and with no specially severe incentives to gamble, like in our main case, but there are also banks, at the top of the overall unconditional distribution of θ_1, which are also heterogeneous in their incentives to gamble (perhaps due to "too-big-to-fail" problems and bail-out expectations). Then, in addition to having a tax on short-term funding that graduates the contribution to systemic externalities of the banks in the non-gambling group (and also, though imperfectly, the banks in the gambling group), it might be socially valuable to introduce a complementary quantitative limit on short-term funding (say, via a stable funding requirement) so as to further control the externalities caused by the group of gambling banks.

9. Dealing with Systemic Importance

Suppose that factors such as size, interconnectedness, lack of substitutability, global (cross-jurisdictional) activity, and complexity make some banks more "systemically important" than others in the very sense that the per-unit contribution of their short-term funding to the systemic risk measure X is larger than for other banks. Suppose in particular that systemic importance is captured by a new dimension of heterogeneity, θ_3, which only enters significantly into the equations of the economy through the following extended measure of systemic risk:

$$X = \int_0^1 \int_0^1 w(\theta_3) x(\theta_1, \theta_3) f(\theta_1, \theta_3) d\theta_1 d\theta_3, \tag{29}$$

where $w(\theta_3)$ is the systemic risk weight of the banks of class θ_3 and $x(\theta_1, \theta_3)$ denotes the short-term funding used by banks characterized by the pair (θ_1, θ_3). This quantity is written as a function θ_3 to account for the possibility that regulation is made contingent on θ_3 and, through it, systemic importance has an impact on the short-term funding used by each bank.

Extending our characterization of competitive equilibria (unregulated or with taxes) and the socially optimal allocation to deal with this case is immediate. Moreover, it can be shown that decentralizing the socially optimal allocation as a competitive equilibrium with taxes will only require setting $\tau(\theta_3) = w(\theta_3)\tau^*$, where $\tau^* = E_z(\varepsilon(x^*(z), z))c'(X^*)$ is a reference rate set exactly like in (12), except z should now be interpreted as the vector (z_1, z_3) of individual bank characteristics. So the presence of heterogenous systemic importance calls for considering each bank's systemic importance measure $w(\theta_3)$ as a scaling factor for the reference (flat) tax rate τ^*.[25] Importantly, the optimal Pigovian tax rate $\tau(\theta_3)$ preserves the key property of being not directly dependent on the individual value of each bank's lending opportunities as measured by θ_1.

10. Conclusions

We have developed a formal analysis of the relative performance of realistic price-based and quantity-based approaches to the regulation of systemic externalities associated with banks' short-term funding. The analysis suggests that if the return to the lending (or investment) activities undertaken by the banks using this funding is heterogeneously distributed across banks (or, similarly, over time), a Pigovian tax on short-term funding will dominate a net stable funding ratio or a liquidity coverage ratio. If some (poorly capitalized or low charter value) banks have strong gambling incentives and expand their activity as a way to shift risk to outside stakeholders (e.g., the deposit insurer), quantity requirements may have better properties. In general terms, an optimal regulatory design may combine price- and quantity-based instruments, and the emphasis on each of them will depend on what is the dominant dimension of heterogeneity across banks (or variation over time).

Going beyond the pure regulation of short-term funding, capital requirements—the most important regulatory instrument in banking—can be seen as a way to directly influence gambling incentives (and the extent to which banks differ in this dimension).

[25] On June 25, 2011 the BCBS issued a press release announcing progress on a consultative document containing a methodology for assessing systemic importance.

Strengthening capital requirements, by ensuring shareholders internalize a larger part of the lower tail of the returns generated by the banks, will tend to produce a smaller measure of banks with strong inclinations for risk shifting. In this sense, our results suggest that in a scenario with stronger capital regulation, such as that envisaged after the full implementation of Basel III, there will be even greater room for having a Pigovian-style tax, levy, or charge on short-term funding as part of the second-best regulatory mix.

References

Acharya, G., D. Gale, and T. Yorulmazer. 2011. "Rollover Risk and Market Freezes." *Journal of Finance* 66 (4): 1177–1209.

Acharya, V., and O. Merrouche. 2010. "Precautionary Hoarding of Liquidity and Inter-Bank Markets: Evidence from the Sub-prime Crisis." NBER Working Paper No. 16395.

Acharya, V., and S. Öncü. 2010. "The Repurchase Agreement (Repo) Market." In *Regulating Wall Street: The Dodd-Frank Act and the New Architecture of Global Finance*, ed. V. Acharya, T. Cooley, M. Richardson, and I. Walters, 319–50 (Chapter 11). Wiley.

Acharya, V., L. Pedersen, T. Philippon, and M. Richardson. 2010. "Measuring Systemic Risk." Federal Reserve Bank of Cleveland Working Paper No. 10-02.

Acharya, V., and S. Viswanathan. 2011. "Leverage, Moral Hazard, and Liquidity." *Journal of Finance* 66 (1): 99–138.

Adrian, T., and M. Brunnermeier. 2008. "CoVaR." Federal Reserve Bank of New York Staff Report No. 348.

Allen, F., A. Babus, and E. Carletti. 2010. "Financial Connections and Systemic Risk." NBER Working Paper No. 16177.

Allen, F., and D. Gale. 2000. "Financial Contagion." *Journal of Political Economy* 108 (1): 1–33.

Basel Committee on Banking Supervision. 2010. "Basel III: International Framework for Liquidity Risk Measurement, Standards and Monitoring." Bank for International Settlements (December).

Brunnermeier, M. 2009. "Deciphering the Liquidity and Credit Crunch 2007–2008." *Journal of Economic Perspectives* 23 (1): 77–100.

Brunnermeier, M., and M. Oehmke. Forthcoming. "The Maturity Rat Race." Forthcoming in *Journal of Finance.*

Calomiris, C., and C. Kahn. 1991. "The Role of Demandable Debt in Structuring Optimal Banking Arrangements." *American Economic Review* 81 (3): 497–513.

Diamond, D., and P. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (3): 401–19.

Diamond, D., and R. Rajan. 2001. "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking." *Journal of Political Economy* 109 (2): 287–327.

Farhi, E., and J. Tirole. 2010. "Collective Moral Hazard, Maturity Mismatch and Systemic Bailouts." Forthcoming in *American Economic Review.*

Financial Services Authority. 2009. *Strengthening Liquidity Standards.* Policy Statement No. 09/16 (October).

Glaeser, E., and A. Shleifer. 2001. "A Reason for Quantity Regulation." *American Economic Review: Papers and Proceedings* 91 (2): 431–35.

Gorton, G. 2009. "Slapped in the Face by the Invisible Hand: Banking and the Panic of 2007." Paper prepared for the Federal Reserve Bank of Atlanta's 2009 Financial Markets Conference: Financial Innovation and Crisis, Jekyll Island, Georgia, May 11–13.

Gorton, G., and A. Metrick. 2010. "Regulating the Shadow Banking System." *Brookings Papers on Economic Activity* 41 (2): 261–312.

Hellmann, T., K. Murdock, and J. Stiglitz. 2000. "Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?" *American Economic Review* 90 (1): 147–65.

Holmstrom, B., and J. Tirole. 1997. "Financial Intermediation, Loanable Funds, and the Real Sector." *Quarterly Journal of Economics* 112 (3): 663–91.

Huang, R., and L. Ratnovski. 2011. "The Dark Side of Bank Wholesale Funding." *Journal of Financial Intermediation* 20 (2): 248–63.

Huberman, G., and R. Repullo. 2010. "Moral Hazard and Debt Maturity." Mimeo, CEMFI.

Jeanne, O., and A. Korinek. 2010. "Managing Credit Booms and Busts: A Pigouvian Taxation Approach." NBER Working Paper No. 16377.

Jensen, M., and W. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3 (4): 305–60.

Kaplow, L., and S. Shavell. 2002. "On the Superiority of Corrective Taxes to Quantity Regulation." *American Law and Economics Review* 4 (1): 1–17.

Keeley, M. C. 1990. "Deposit Insurance, Risk, and Market Power in Banking." *American Economic Review* 80 (5): 1183–1200.

Keeton, W. 1990. "Risk-Based Deposit Insurance: Is Price Regulation Necessarily Better than Quantity Regulation?" In *Game Plans for the '90s*, 44–61. Proceedings of the 26th Annual Conference on Bank Structure and Competition, Federal Reserve Bank of Chicago, May 9–11.

Laeven, L., and F. Valencia. 2010. "Resolution of Banking Crises: The Good, the Bad, and the Ugly." IMF Working Paper No. 10/146.

Lorenzoni, G. 2008. "Inefficient Credit Booms." *Review of Economic Studies* 75 (3): 809–33.

Martin, A., D. Skeie, and E. L. von Thadden. 2010. "Repo Runs." Mimeo, University of Mannheim.

Perotti, E., and J. Suarez. 2009. "Liquidity Insurance for Systemic Crises." CEPR Policy Insight No. 31 (February).

Poole, W. 1970. "Optimal Choice of Monetary Policy Instruments in a Simple Stochastic Macro Model." *Quarterly Journal of Economics* 84 (2): 197–216.

Repullo, R. 2004. "Capital Requirements, Market Power, and Risk-Taking in Banking." *Journal of Financial Intermediation* 13 (2): 156–82.

Rochet, J.-C., and X. Vives. 2004. "Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?" *Journal of the European Economic Association* 2 (6): 1116–47.

Segura, A., and J. Suarez. 2011. "Dynamic Maturity Transformation." Mimeo, CEMFI.

Wagner, W. 2010. "Diversification at Financial Institutions and Systemic Crises." *Journal of Financial Intermediation* 19 (3): 373–86.

Weitzman, M. L. 1974. "Prices vs. Quantities." *Review of Economic Studies* 41 (4): 477–91.

Discussion of "A Pigovian Approach to Liquidity Regulation"

Ernst-Ludwig von Thadden
University of Mannheim

The regulation of bank liquidity has been one of the most controversial topics in the recent debate about the reform of financial markets. The new regulatory framework for banks published by the Basel Committee on Banking Supervision on December 16, 2010, enters new regulatory territory on several fronts. But arguably the most daring and novel rules concern bank liquidity. The "net stable funding ratio," a requirement to be introduced until 2018, and the "liquidity coverage ratio," for implementation in 2015, are regulatory attempts to strengthen bank balance sheets against liquidity shocks, introduced under the impression that the sudden collapses of banks in the Great Financial Crisis were avoidable and systemically dangerous.

This impression is quite certainly correct, but there has been very little research in finance supporting the implementation of such instruments. The difficulties of formulating a theory of the systemic role of bank balance sheets are indeed considerable, ranging from the question of the transmission of idiosyncratic shocks in interbank markets to the questions of the costs and benefits of maturity transformation and the inefficiency of markets with cash-in-the-market pricing. In their paper (in this issue), Perotti and Suarez cut through all this maze and present a simple and powerful analysis of a basic externality underlying the regulation of bank liquidity.

They do this in the simplest possible model of banking regulation: a static, partial-equilibrium, one-good model of competing firms with one aggregate externality. This approach makes it possible to exhibit the market failure and the regulatory remedy very clearly. The simplicity of this approach is the strength of the paper and its greatest weakness at the same time, because the model is tailored to understand exactly the question it is supposed to highlight. This is very economical, but it makes it difficult to evaluate how other aspects of banking and financial markets interfere with banking liquidity and why the regulation of banking liquidity may, after all, be such a complicated and controversial problem.

To see the argument, one can consider the following model (which simplifies some of the math in the paper, such as replacing integrals with sums, without changing anything of the substance). There are n banks, with weights p_i that reflect size or systemic importance. Bank i creates an expected net present value (NPV) of $v_i(x_i, X)$, where x_i is short-term borrowing by bank i and $X = \sum_{i=1}^{n} p_i x_i$ is aggregate short-term borrowing. The crucial assumptions are that

$$\frac{\partial v_i}{\partial X} < 0 \text{ and } \frac{\partial^2 v_i}{\partial x_i \partial X} < 0, \tag{1}$$

i.e., that aggregate short-term borrowing reduces expected individual bank value overall and at the margin. The main specification given for this reduced form is

$$v_i(x_i, X) = \pi_i(x_i) - \varepsilon_i(x_i)c(X), \tag{2}$$

where π_i is bank value in the absence of crises, ε_i the probability that bank i is affected by a crisis, and c the expected cost of a crisis for a single bank.

In a competitive equilibrium, each bank chooses its short-term borrowing x_i^e to maximize its expected NPV, taking aggregate borrowing X^e as given. Under the appropriate concavity and boundary assumptions, this yields the first-order conditions

$$\frac{\partial}{\partial x_i} v_i\left(x_i^e, X^e\right) = 0 \tag{3}$$

that characterize the equilibrium.

A social planner, on the other hand, will choose borrowing levels x_i^* that take the aggregate externality into account, which yields the first-order conditions

$$\frac{\partial}{\partial x_i} v_i\left(x_i^*, X^*\right) + \sum_j p_j \frac{\partial}{\partial X} v_j\left(x_j^*, X^*\right) = 0. \tag{4}$$

By assumption 1, these conditions imply that $X^* < X^e$, hence that the competitive equilibrium involves too much short-term borrowing.

As in basic public finance, it is then simple to see that a linear (Pigovian) tax on short-term borrowing, to the tune of

$$\tau^* = -\sum_j p_j \frac{\partial}{\partial X} v_j\big(x_j^*, X^*\big), \tag{5}$$

perfectly aligns public and private incentives in equilibrium and thus achieves the socially optimal level of short-term borrowing at each bank. Furthermore, it can be seen easily that a quantity regulation of the sort

$$x_i \leq \bar{x}, \tag{6}$$

while generating the optimal aggregate borrowing level X^*, cannot restore first-best efficiency, because it would shift short-term borrowing from banks with a high marginal product to those with a low marginal product.

This brief account is a fairly complete summary of two-thirds of the analysis in the paper. Much of the value of the paper lies in the remaining third of the analysis, which applies the above simple model to several issues that are of central concern to bank regulators.

First, the authors discuss the "liquidity coverage ratio" of Basel III. This requirement can be modeled as requiring banks to hold a certain fraction ϕ of their short-term borrowing in liquid assets instead of investing it in (hopefully) positive-NPV projects. Holding an amount m_i in such liquid assets involves a cost of $\delta_i m_i$, which is the difference between the bank's borrowing rate and the safe rate at which it can hold the liquid asset. The bank's objective function then becomes

$$v_i(x_i - m_i, \hat{X}) - \delta_i m_i, \tag{7}$$

where $\hat{X} = \sum_{i=1}^{n} p_i(x_i - m_i)$. The bank maximizes (7) under the constraint $\phi x_i \leq m_i$, taking $\hat{X}$ as given. Although the liquidity coverage requirement is formally similar to imposing a tax rate $\tau_i = \frac{\delta_i \phi}{1-\phi}$, it is not difficult to see that the resulting equilibrium allocation is not first-best, even if the induced tax rate is set to the first-best level τ_i^*, because it involves a deadweight loss of $\tau_i^* X^*$ (proposition 6).

The second extension of the paper is a simple model of moral hazard. Moral hazard is modeled as an exogenous propensity to borrow at the expense of outside stakeholders, such as the deposit insurer. This propensity is measured by a parameter θ that modifies the special objective function (2) as follows:

$$v_i(x_i, X, \theta) = \pi_i(x_i) - (1 - \theta)\varepsilon_i(x_i)c(X). \tag{8}$$

Hence, a type θ bank ignores a fraction θ of the expected losses it generates. If there is unobservable heterogeneity in θ, a Pigovian tax no longer achieves first-best, because the tax rate would have to condition on θ. However, if θ is the only source of unobserved heterogeneity—that is, if all v_i are identical—then the quantity restriction (6) achieves the first-best. This is straightforward, because now the first-best borrowing amount x_i^{**} is the same for all banks, which the regulator can simply impose as an upper limit (that will of course be attained).

The baseline model and the second extension therefore represent two extremes in which two extreme forms of regulation are optimal. In the baseline model, a price regulation (via linear taxes) is optimal; in the second extension, a quantity regulation is optimal. The optimal regulation in the general case, with unobserved heterogeneity in bank profitability and bank propensity to moral hazard, is an open issue. As the authors conjecture in the introduction, it may follow a similar trade-off as the classic analysis of Weitzman (1974).

The model is simple and exhibits an important externality in banking very clearly. Yet, by focusing on short-term borrowing as the sole source of bank funding, it oversimplifies and makes it difficult to interpret the results in terms of regulatory recommendations. Clearly, banks are funded through a variety of liabilities, not just short-term debt. Incidentally, deposits, which are a classical form of very short-term borrowing, have proven to be a highly reliable source of funding in the recent crisis. But more importantly, bank value is created also by long-term borrowing or equity, two elements that are also absent from the reduced-form profit functions (2) and (8). In these formulations, any interference with short-term borrowing x_i automatically impacts the bank's asset side, because there is nothing else to fund the assets. Regulation in practice, of course, does not necessarily aim at the asset side, but rather tries to make sure that assets are funded through a more resilient liability structure.

To see this issue more clearly, consider the following simplified bank balance sheet:

		d	deposits
m	liquid assets	x	short-term market funding
y	productive assets	b	long-term borrowing
		e	equity

A regulation such as the net stable funding ratio of Basel III does not primarily intend to influence y, but targets the bank's potential maturity mismatch by relating b and e to y. Short-term borrowing x may be used by the bank as an instrument to achieve a net stable funding objective, but x is not mechanically targeted by this type of regulation, as the present paper assumes.

This immediately raises the next question: why do banks use short-term funding and what types of short-term funding are used in what way in the value-creation process? Clearly, maturity transformation is the central feature of short-term borrowing and of significant economic benefit. As argued by Diamond and Dybvig (1983) for the case of commercial banking and by Martin, Skeie, and von Thadden (2010) for the case of dealers and investment banks, short-term borrowing can be a viable source for long-term value creation. By limiting the use of short-term market finance, regulators can improve the stability of banks (Martin, Skeie, and von Thadden 2010), but this will have consequences for the amount of maturity transformation banks can provide. Is there a stability-profitability trade-off? The present paper yields an unambiguously positive answer to this question, because the trade-off is hard-wired into the objective function (2). It would be interesting to see whether such an assumption can be microfounded by a more detailed model.

Similarly, the assumption about the impact of aggregate borrowing in (1) is a rather extreme one. According to this assumption, aggregate borrowing reduces expected bank profitability and marginal expected profitability at the bank level. This assumption is not very plausible in "good" times (which is in line with the authors' thinking), but whether the possible negative impact in "bad" times is as dominant as the authors assume depends on the functioning or malfunctioning of financial markets under stress. A key problem in the 2008 crisis, as in the looming European banking crisis of 2011, has been the failure of the interbank market. Is this failure simply a consequence of too much short-term borrowing? Or is it possible that safeguards or interventions on the interbank market (such as transparency requirements or liquidity assistances) can remedy such imperfections more efficiently than liquidity constraints on banks? Is it rather secured interbank lending or unsecured interbank lending that has the potential to destabilize the market (Heider and Hoerova 2009)? In this perspective it is rather unfortunate that the paper almost completely abstracts from prices. Even the analysis of competitive equilibrium works only with equilibrium quantities and describes profits in reduced form. This suggests that either asset supply is completely elastic or that the analysis is only partial equilibrium and thus ignores potential feedback effects. Both of these alternatives are not entirely convincing.

While I believe that the thrust of the argument that the authors formulate so simply and elegantly is correct, I would like to see a more detailed description and analysis of the underlying frictions that yield the reduced form that they present. In this sense, the present paper is rather a first step to our understanding than a final treatment of the complex issues that make short-term finance a systemic problem in modern financial markets.

References

Basel Committee on Banking Supervision. 2010. “Basel III: International Framework for Liquidity Risk Measurement, Standards, and Monitoring.” Bank for International Settlements (December). Available at http://www.bis.org/publ/bcbs188.pdf

Diamond, D., and P. Dybvig. 1983. “Bank Runs, Deposit Insurance and Liquidity.” *Journal of Political Economy* 91 (3): 401–19.

Heider, F., and M. Hoerova. 2009. “Interbank Lending, Credit Risk Premia, and Collateral.” *International Journal of Central Banking* 5 (4): 5–43.

Martin, A., D. Skeie, and E.-L. von Thadden. 2010. “Repo Runs.” Mimeo, Federal Reserve Bank of New York.

Weitzman, M. L. 1974. “Prices vs. Quantities.” *Review of Economic Studies* 41 (4): 477–91.

Macroeconomic Propagation under Different Regulatory Regimes: Evidence from an Estimated DSGE Model for the Euro Area*

Matthieu Darracq Pariès, Christoffer Kok Sørensen, and Diego Rodriguez-Palenzuela
European Central Bank

This article analyzes the role of credit market frictions in business-cycle fluctuations and in the transmission of monetary policy. We estimate a closed-economy dynamic stochastic general equilibrium (DSGE) model for the euro area with financially constrained households and firms and embedding an oligopolistic banking sector facing capital constraints. Using this setup we examine the monetary policy implications of the various financial frictions to credit supply and demand and furthermore examine the real economic implications of increasing capital requirements and of introducing risk-sensitive capital requirements. Moreover, the potential for introducing countercyclical bank capital rules and aligning macroprudential tools with standard monetary policy tools is examined. In particular, the model results highlight the importance of operating with a protracted implementation schedule of new regulatory requirements for smoothing out the transitional costs to the economy arising from a more capital-constrained banking sector.

JEL Codes: E4, E5, F4.

*We thank Bernard-Daniel Solomon for valuable input as well as Ignazio Angeloni, Andrew Powell, Frank Smets, and participants and discussants at various conferences, workshops, and seminars held at the ECB, CEIS Tor Vergata, and the IJCB 2011 Spring Conference for stimulating and helpful comments. The views expressed are solely our own and do not necessarily reflect those of the European Central Bank. Corresponding author e-mail: matthieu.darracq_paries@ecb.int.

1. Introduction

The financial crisis which started in 2007 brought to the fore the importance of the financial sector and its potential amplifying effects on business-cycle fluctuations. The massive write-downs and losses that banks had to incur over this period significantly impaired their liquidity and capital positions, which in turn forced many banks to cut back on activities and to shed assets. This deleveraging process in the banking sector may have hampered the access to financing for some bank-dependent borrowers and thereby reduced their ability to consume and invest, potentially reinforcing the economic downturn. Whereas in the macroeconomic literature it has long been recognized that financial intermediation may play a role in economic fluctuations through the financial accelerator mechanism relating to the banks' borrowers,[1] the possible amplifying impact on the business cycle of shocks directly hitting the financial intermediaries has only recently been taken up by the literature.[2] The importance of the banks' balance sheet situation in transmitting shocks to monetary policy (and other types of shocks) has, however, long been recognized in the empirical literature. For example, it has been pointed out that more liquid and well-capitalized banks are better able to absorb shocks hitting the macroeconomic environment (including changes in monetary policy) than more capital- and liquidity-constrained

[1] Financing frictions arising in the context of asymmetric information between borrowers and lenders are often suggested as a prime candidate for endogenously amplifying and increasing the persistence of even small transitory exogenous shocks. The basic idea, often called the financial accelerator, is that in the presence of credit constraints exogenous shocks can generate a positive feedback effect between the financial health of borrowing firms or households and output; see, e.g., Carlstrom and Fuerst (1997), Kiyotaki and Moore (1997), and Bernanke, Gertler, and Gilchrist (1999) (BGG hereafter). Recent work by Christiano, Motto, and Rostagno (2007), Christensen and Dib (2008), and Liu, Wang, and Zha (2009) quantifies the interlinkages between the financial and real sectors using a financial accelerator mechanism.

[2] For some recent studies modeling the banking sector in a DSGE modeling framework, see, e.g., Meh and Moran (2008), Van den Heuvel (2008), Agenor and Alper (2009), Agenor and Pereira da Silva (2009), Aguiar and Drumond (2009), Angeloni and Faia (2009), Covas and Fujita (2009), de Walque, Pierrard, and Rouabah (2009), Dib (2009), Gertler and Karadi (2009), Christiano, Motto, and Rostagno (2010), and Gerali et al. (2010).

banks.[3] In addition to the attention on the role of financial intermediaries brought forward by the financial crisis, the introduction of more risk-sensitive capital requirements (i.e., the Basel II capital adequacy framework; see Basel Committee on Banking Supervision 2006) has reinforced the concerns that financial intermediation by itself might have substantial feedback effects on the real economy. Moreover, as a consequence of the financial crisis, at the end of 2010 the Basel Committee on Banking Supervision (BCBS) introduced amendments to the bank regulatory framework (i.e., Basel III) with the aim of strengthening capital requirements. The new requirements will be gradually phased in as of 2013 and are scheduled to be fully implemented by 2019.[4] Our model is also well suited for analyzing the potential costs (and benefits) of moving towards higher capital ratio targets and the role of monetary policy during such a transition. Furthermore, the financial crisis has reinforced interest in macroprudential tools and policies that might be applied by policymakers to reduce the risks of financial boom and bust cycles and thereby lead to a more stable path of real economic growth. A general equilibrium framework, such as ours, is also useful for analyzing the potential for macroprudential tools and their interaction with other macroeconomic and monetary policy instruments.

Against this background, in this article we propose a closed-economy DSGE model with financial frictions including a banking sector which faces monopolistic competition and is subject to capital constraints. The latter may owe both to market disciplining forces (i.e., banks operate with a capital buffer) and to regulatory capital adequacy rules (which can be either risk insensitive or risk sensitive). Furthermore, the presence of monopolistic competition in the banking sector gives rise to some degree of stickiness in banks' adjustment of lending and deposit rates to changes in monetary policy rates. From a theoretical viewpoint, a sluggish pass-through of bank loan and deposit rates to policy rate changes is based on the notion of banks having some degree of market power, which may derive from banks being "special" in the sense of being able to reduce (by acting

[3]See, e.g., Bernanke and Lown (1991), Peek and Rosengren (1995), Kashyap and Stein (2000), Van den Heuvel (2002), Gambacorta and Mistrulli (2004), and Kishan and Opiela (2006).

[4]See Basel Committee on Banking Supervision (2011).

as "delegated monitors") the information gap between savers and borrowers of funds.[5] In general, banks' interest rate setting behavior can be expected to depend on the degree of bank competition (or market power of banks) and on factors related to the costs of financial intermediation (such as interest rate and credit risk, menu costs and other operational costs, banks' degree of risk aversion, and the cost of non-deposit funding sources).[6] Hence, by exploiting their market power, banks are able to generate profits and thus to replenish their capital buffers following shocks to their liquidity and capital positions.[7] Under risk-sensitive capital requirements, banks' capital positions are affected by changes in the risk profile of their borrowers over the business cycle, and the time-varying nature of bank borrower risk profiles is therefore also considered in our modeling of firms and households.

On the real side of the economy, we assume that households and firms are financially constrained in their spending and investment decisions, and we furthermore incorporate some degree of heterogeneity in the household sector. The model has a subset of firms that are financially constrained and can only borrow by using revenue and capital as collateral, and a subset of financially constrained households that use debt collateralized by housing and part of their wage income. Both firms and households are affected by idiosyncratic shocks to their collateral values. Firms and households default on their loans when the value of their collateral is below the repayment promised to the lender. In order to keep the model tractable, we follow other DSGE models of financial frictions in using differences in the level of impatience of agents to generate equilibrium borrowing and lending (e.g., Iacoviello 2005). In equilibrium, more impatient agents (borrowers and entrepreneurs) will borrow from patient savers.

[5]See, e.g., Diamond and Dybvig (1983), Diamond (1984), and Diamond and Rajan (2001).

[6]There is ample empirical evidence for the existence of a sluggish bank interest rate pass-through in the euro area (see, e.g., Mojon 2001, de Bondt 2005, Sander and Kleimeier 2006, Kok Sørensen and Werner 2006, and Gropp, Kok Sørensen, and Lichtenberger 2007).

[7]There are a few recent studies that embed features of an incomplete bank interest rate pass-through into a DSGE model framework; see, e.g., Kobayashi (2008), Agenor and Alper (2009), Hülsewig, Mayer, and Wollmershäuser (2009), and Gerali et al. (2010).

Building on Notarpietro (2007) and Iacoviello and Neri (2010), we define a three-agent, two-sector economy, where the impatient agents face collateral requirements when asking for mortgages or loans. Firms produce non-durable consumption goods and residential goods. The latter serve two purposes: they can be directly consumed, thus providing utility services as any durable good, or they can be used as collateral in the credit market, to obtain extra funds for financing consumption. The role of collateral constraints in closed economies has been estimated in DSGE models by Notarpietro (2007) and Iacoviello and Neri (2010), who report the relevance of housing market shocks in shaping consumption dynamics in the United States. Most existing models of household borrowing in a DSGE framework follow Kiyotaki and Moore (1997) and Iacoviello (2005) in using a hard borrowing constraint and assuming it always binds. The Kiyotaki-Moore model of credit constraints can be seen as a special case of the current model in which there is no uncertainty about the future value of the collateral when the loan is made. The assumption that the constraint always binds makes the leverage ratio in their model constant. Furthermore, they ignore any difference between borrowing rates and the risk-free rate. The model proposed here can at least qualitatively match the typically observed countercyclical leverage ratio of households.[8] The assumption of an always binding borrowing constraint is questionable for large shocks that may be of particular interest to policymakers, and it may severely distort the dynamics of borrowers and the rest of the economy in those circumstances. The soft borrowing constraint in our model (with interest rates rising smoothly as a function of borrowing) will always bind as long as it can be satisfied.

The only other papers that have allowed for financing frictions affecting both households and firms are Iacoviello (2005) and Gerali et al. (2010). Both of these papers rely on hard borrowing constraints, as in Kiyotaki and Moore (1997), to model credit frictions and assume the borrowing constraints always bind. Our model setup provides an alternative perspective by including costs of default and positive lending spreads.

By allowing for frictions concerning both credit demand and supply, the contributions of this paper cover several dimensions. First,

[8]For instance, as found for the United States by Adrian and Shin (2009).

apart from encompassing the traditional financial accelerator mechanism arising in the context of financially constrained borrowers, our model allows for assessing the impact of frictions within the banking sector, such as its price-setting behavior and constraints to its capital management. In particular, we assess the extent to which the presence of bank loan and deposit rate sluggishness affect monetary policy optimization. Moreover, our setup allows for examining the macroeconomic implications of shocks to bank capital (such as those observed during the 2007–10 financial crisis as well as reflected in the proposal to introduce stronger capital requirements under the Basel III agreement) and the implications of introducing risk-sensitive capital requirements or the transitional effects of higher capital requirements. Furthermore, our model can also shed some light on the potential effects of active macroprudential policies over the cycle and their interaction with monetary policy.

At the same time, our current model setup is less suited for analyzing the issues of liquidity and wholesale funding vulnerabilities, which arguably were other main contributing factors to the severity and propagation of the financial crisis. The macroeconomic implications of money market disruptions and the potential role of unconventional monetary policies have been addressed in other recent papers (see, e.g., Gertler and Kiyotaki 2009 and also Christiano, Motto, and Rostagno 2010).

The rest of the article is organized as follows. Section 2 describes the main decision problems of the structural model.[9] Section 3 presents the results of the Bayesian estimation, while section 4 explores the business-cycle implications of the financial frictions; in particular, the optimal monetary policy responses under different regulatory frameworks are investigated, focusing on the introduction of higher and risk-based capital requirements and macroprudential rules. Section 5 concludes.

2. Theoretical Model

The real side of the economy is modeled as a three-agent, two-sector economy, producing residential and non-residential goods.

[9]For the purpose of brevity, many model details can be found in the working paper version of this article, Darracq Pariès, Kok Sørensen, and Rodriguez-Palenzuela (2010).

Residential goods are treated here as *durable* goods. A continuum of entrepreneurs, with unit mass, produce non-residential and residential intermediate goods under perfect competition and face financing constraints. Retailers differentiate the intermediate goods under imperfect competition and staggered price setting, while competitive distribution sectors serve final non-residential consumption as well as residential and non-residential investments. A continuum of infinitely lived households, with unit mass, is composed of two types, differing in their relative intertemporal discount factor. A fraction $(1 - \omega)$ of households are relatively *patient*, the remaining fraction ω being *impatient*. Households receive utility from consuming both non-residential and residential goods, and disutility from labor. Impatient households are financially constrained. The labor market structure is characterized by homogeneous labor supply and monopolistically competitive unions, which gives rise to staggered wage setting.

The banking sector collects deposits from patient households and provides funds to entrepreneurs and impatient households. Three layers of frictions affect financial intermediaries. First, wholesale bank branches face capital requirements (which can be risk insensitive or risk sensitive) as well as adjustment costs related to their capital structure. Second, some degree of nominal stickiness generates some imperfect pass-through of market rates to bank deposit and lending rates. Finally, due to asymmetric information and monitoring costs in the presence of idiosyncratic shocks, the credit contracts proposed to entrepreneurs and impatient households factor in external financing premia which depend indirectly on the borrower's leverage. Figure 1 provides an overview of the financial contracts linking the banking sector to the real economy.

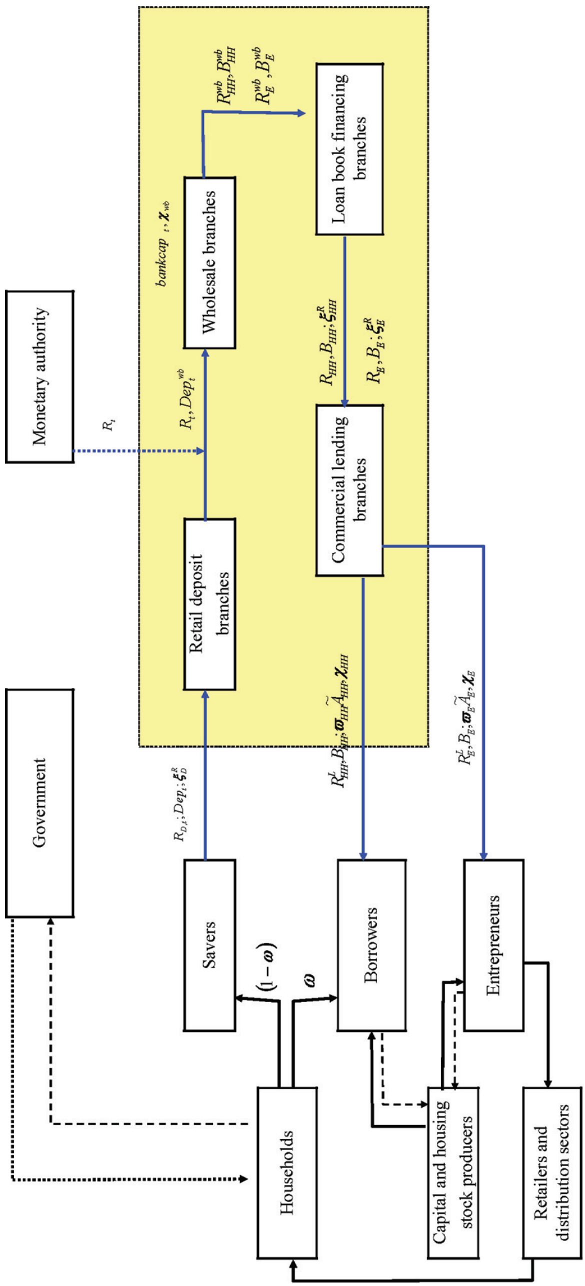
Finally, a government sector collecting taxes and providing lump-sum fiscal transfers and a monetary authority applying a standard Taylor rule close the model.

2.1 Households

2.1.1 The Saver's Program

The patient agents, $s \in [\omega, 1]$, are characterized by a higher intertemporal discount factor than the borrowers, and thus act as net lenders

Figure 1. Structure of the Model



in equilibrium. They own the productive capacities of the economy. Each patient agent receives instantaneous utility from the following instantaneous utility function:

$$\mathcal{W}_t^s = \mathbb{E}_t\left\{\sum_{j\geq 0}\gamma^j\left[\begin{array}{l}\frac{1}{1-\sigma_X}\left(X_{t+j}^s\right)^{1-\sigma_X} - \frac{\varepsilon_{t+j}^L\overline{L}_{S,C}}{1+\sigma_{LC}}\left(N_{Ct+j}^s\right)^{1+\sigma_{LC}}\\ \qquad -\frac{\varepsilon_{t+j}^L\overline{L}_{S,D}}{1+\sigma_{LD}}\left(N_{Dt+j}^s\right)^{1+\sigma_{LD}}\end{array}\right]\varepsilon_{t+j}^{\beta}\right\},$$

where X_t^s is an index of consumption services derived from non-residential final goods (C^s) and residential stock (D^s), respectively.

$$X_t^s \equiv \left[\left(1-\varepsilon_t^D\omega_D\right)^{\frac{1}{\eta_D}}\left(C_t^s - h_S C_{t-1}^s\right)^{\frac{\eta_D-1}{\eta_D}} + \varepsilon_t^D\omega_D^{\frac{1}{\eta_D}}\left(D_t^s\right)^{\frac{\eta_D-1}{\eta_D}}\right]^{\frac{\eta_D}{\eta_D-1}},$$

with the parameter h_S capturing habit formation in consumption of non-residential goods. We introduce three stochastic terms in the utility function: a preference shock ε_t^{β}, a labor supply shock ε_t^L (common across sectors), and a housing preference shock ε_t^D. The latter affects the relative share of residential stock, ω_D, and modifies the marginal rate of substitution between non-residential and residential goods consumption. All the shocks are assumed to follow stationary AR(1) processes.

Households receive disutility from their supply of homogenous labor services to each sector, $N_{C,t}^s$ and $N_{D,t}^s$. The real compensation of hours worked in each sector are denoted $w_{C,t}^s$ and $w_{D,t}^s$. The specification of labor supply assumes that households have preferences over providing labor services across different sectors. In particular, the specific functional form adopted implies that hours worked are perfectly substitutable across sectors. $\overline{L}_C$ and $\overline{L}_D$ are level-shift terms needed to ensure that the patient's labor supply is equal to one in the steady state.

The saver maximizes its utility function subject to an infinite sequence of the following budget constraint:

$$\begin{aligned} &C_t^s + Q_{D,t}T_{D,t}\left(D_t^s - (1-\delta)D_{t-1}^s\right) + Dep_t^s \\ &\quad = \frac{(1+R_{D,t-1})}{(1+\pi_t)}Dep_{t-1}^s + (1-\tau_{w,t})\left(w_{C,t}^s N_{C,t}^s + w_{D,t}^s N_{D,t}^s\right) \\ &\qquad + \Pi_t^s + TT_t^s, \end{aligned}$$

where $Q_{D,t}T_{D,t}$ is real price of housing stock in terms of non-residential goods, TT_t^s are real government transfers, and Π_t^s are real distributed profits. $\delta \in (0,1)$ is the residential good depreciation rate. π_t is the non-residential good inflation rate. $R_{D,t-1}$ is the nominal interest rate paid on the one-period real deposits Dep_t^s.

In equilibrium, all savers have identical consumption plans. Therefore, we can drop the superscript s. We also allow for a time-varying labor income tax, given by $1 - \tau_{w,t} = (1 - \overline{\overline{\tau}}_w)\varepsilon_t^W$.

2.1.2 The Borrower's Program[10]

Each impatient agent $b \in [0, \omega]$ receives utility from the same type of function as in the case of patient households but with a lower discount factor $\beta < \gamma$:[11]

$$\mathcal{W}_t^b = \mathbb{E}_t \left\{ \sum_{j \geq 0} \beta^j \begin{bmatrix} \frac{1}{1-\sigma_X} \left(\widetilde{X}_{t+j}^b\right)^{1-\sigma_X} - \frac{\varepsilon_{t+j}^L \overline{L}_{B,C}}{1+\sigma_{LC}} \left(N_{C,t+j}^b\right)^{1+\sigma_{LC}} \\ - \frac{\varepsilon_{t+j}^L \overline{L}_{B,D}}{1+\sigma_{LD}} \left(N_{D,t+j}^b\right)^{1+\sigma_{LD}} \end{bmatrix} \varepsilon_{t+j}^\beta \right\},$$

where $\widetilde{X}_t^b$ is given by

$$\widetilde{X}_t^b \equiv \left[\left(1 - \varepsilon_t^D \omega_D\right)^{\frac{1}{\eta_D}} \left(\widetilde{C}_t^b - h_B \widetilde{C}_{t-1}^b\right)^{\frac{\eta_D - 1}{\eta_D}} + \varepsilon_t^D \omega_D^{\frac{1}{\eta_D}} \left(\widetilde{D}_t^b\right)^{\frac{\eta_D - 1}{\eta_D}} \right]^{\frac{\eta_D}{\eta_D - 1}}.$$

As regards savers, $\overline{L}_{B,C}$ and $\overline{L}_{B,D}$ are level-shift terms needed to ensure that the impatient's labor supply equals one in the steady state.

Borrowers' incomes and housing stock values are subject to common idiosyncratic shocks $\varpi_{HH,t}$ that are i.i.d. across borrowers and across time. $\varpi_{HH,t}$ has a log-normal cumulative distribution function (CDF) $F(\varpi)$ with $F'(\varpi) = f(\varpi)$, and a mean of $E(\varpi) = 1$. The variance of the idiosyncratic shock $\sigma_{HH,t}$ is time varying. The value of the borrower's house is given by

$$\varpi_{HH,t} \widetilde{Q}_{D,t} T_{D,t} (1 - \delta) \widetilde{D}_{t-1}^b.$$

---

[10]The specification adopted here is broadly similar to Solomon (2011).

[11]Variables related to the saver are denoted with a superscript b, as opposed to s, used for the savers.

Lending in this economy is only possible through one-period state-contingent debt contracts that require a constant repayment of $\frac{(1+R^L_{HH,t})}{1+\pi_t}B_{HH,t-1}$ independent of $\varpi_{HH,t}$ if the borrower is to avoid costly loan monitoring or enforcement, where $R^L_{HH,t}$ is the nominal lending rate.

The borrower can default and refuse to repay the debt. Savers cannot force borrowers to repay. Instead, lending must be intermediated by commercial banks that have a loan enforcement technology allowing them to seize collateral expressed in real terms

$$\varpi_{HH,t}\tilde{A}^b_{HH,t} = (1-\chi_{HH})\varpi_{HH,t}\widetilde{Q}_{D,t}T_{D,t}(1-\delta)\widetilde{D}^b_{t-1}$$

at a proportional cost $\mu_{HH}\varpi_{HH,t}\tilde{A}_{HH,t}$ when the borrower defaults.

$\mu_{HH} \in (0,1)$ determines the deadweight cost of default; $0 < \chi_{HH} \leq 1$ represents housing exemptions. It defines the maximum loan-to-collateral ratio (often called the loan-to-value ratio) that the bank is willing to grant against each component of the collateral. Conditional on enforcement, the law cannot prevent the bank from seizing $\varpi_{HH,t}\tilde{A}_{HH,t}$. Suppose first that the borrower does not have access to any insurance against the $\varpi_{HH,t}$ shock. Whenever $\varpi_{HH,t} < \overline{\varpi}_{HH,t}$, the borrower prefers to default and lose

$$\varpi_{HH,t}\tilde{A}^b_{HH,t} < \frac{\left(1+R^L_{HH,t}\right)}{1+\pi_t}B_{HH,t-1} = \overline{\varpi}_{HH,t}\tilde{A}^b_{HH,t}$$

when the bank enforces the contract. On the other hand, when $\varpi_{HH,t} \geq \overline{\varpi}_{HH,t}$, the borrower prefers to pay $\frac{(1+R^L_{HH,t})}{1+\pi_t}B_{HH,t-1}$ rather than lose $\varpi_{HH,t}\tilde{A}_{HH,t} \geq \frac{(1+R^L_{HH,t})}{1+\pi_t}B_{HH,t-1}$.

To be able to use a representative agent framework while maintaining the intuition of the default rule above, we assume that borrowers belong to a large family that can pool their assets and diversify away the risk related to $\varpi_{HH,t}$ after loan repayments are made. As in Lucas (1990) and Shi (1997), the family maximizes the expected lifetime utility of borrowers with an equal welfare weight for each borrower. The payments from the insurance scheme cannot be seized by the bank. As a result, despite the insurance, the bank cannot force the borrower to repay $\frac{(1+R^L_{HH,t})}{1+\pi_t}B_{HH,t-1}$ when $\varpi_{HH,t} < \overline{\varpi}_{HH,t}$. Like the individual borrowers, the family cannot commit to always repay the loan (or make up for any lack of payment by a borrower), even though from an ex ante perspective it is optimal to do so. Ex post, from the perspective of maximizing the expected welfare of the borrowers, for any given $R^L_{HH,t}$ it is optimal to have borrowers with $\varpi_{HH,t} < \overline{\varpi}_{HH,t}$ default and borrowers with $\varpi_{HH,t} \geq \overline{\varpi}_{HH,t}$ repay $\frac{(1+R^L_{HH,t})}{1+\pi_t} B_{HH,t-1}$.

Given the large family assumption in particular, households' decisions are the same in equilibrium. Therefore, we can drop the superscript b.

By pooling the borrowers' resources, the representative family has the following aggregate repayments and defaults on its outstanding loan:

$$H(\overline{\varpi}_{HH,t})\tilde{A}_{HH,t} = \left[(1 - F_t(\overline{\varpi}_{HH,t}))\overline{\varpi}_{HH,t} + \int_0^{\overline{\varpi}_{HH,t}} \varpi dF_t\right] \tilde{A}_{HH,t}.$$

On the commercial lending bank side, the profit made on the credit allocation is given by

$$G(\overline{\varpi}_{HH,t})\tilde{A}_{HH,t} - \frac{(1 + R_{HH,t-1})}{1 + \pi_t} B_{HH,t-1} \geq 0$$

with

$$G(\overline{\varpi}_{HH,t}) = (1 - F_t(\overline{\varpi}_{HH,t}))\overline{\varpi}_{HH,t} + (1 - \mu_{HH}) \int_0^{\overline{\varpi}_{HH,t}} \varpi dF_t.$$

$R_{HH,t-1}$ is the interest rate at which the commercial lending bank gets financing every period, while $R^L_{HH,t}$ is the state-contingent lending rate. Competition among banks will ensure that profits are null in equilibrium. The zero-profit condition could also be seen as the borrowing constraint in this model. Notice that this constraint always binds as long as it can be satisfied.[12] In contrast, the hard borrowing constraint in Kiyotaki and Moore (1997) or Iacoviello (2005) may not bind, even though authors using that framework assume it always binds to allow the use of perturbation methods.[13]

[12]If the constraint were slack, the lender could always reduce the borrower's expected repayments while still respecting the constraint by reducing $\overline{\varpi}_{HH,t}$.

[13]This may be a reasonable assumption for small shocks, but it can be a bad approximation for larger shocks that may be of concern to policymakers.

The caveat is that if a new shock significantly lowers the value of $\tilde{A}_{HH,t}$, it may be impossible to find a default threshold that allows the bank to break even on the loan with the risk-free rate. This should not be a major concern except for very low aggregate shock values.[14]

With the assumption of perfectly competitive banks, we can represent the problem of borrowers as if they choose default thresholds as a function of the aggregate states directly, subject to the bank's participation constraints.

Each borrower maximizes utility function with respect to $(\widetilde{C}_t, \widetilde{D}_t, B_{HH,t}\overline{\varpi}_{HH,t}, N_{C,t}, N_{D,t})$ subject to an infinite sequence of real budget constraints:[15]

$$\begin{aligned}\widetilde{C}_t + \widetilde{Q}_{D,t}T_{D,t}(\widetilde{D}_t - (1-\delta)\widetilde{D}_{t-1}) + H(\overline{\varpi}_{HH,t})\tilde{A}_{HH,t} &= B_{HH,t} + \widetilde{TT}_t \\ &+ \widetilde{w}_{C,t}\widetilde{N}_{C,t} + \widetilde{w}_{D,t}\widetilde{N}_{D,t}\end{aligned}$$

and the zero-profit condition for the commercial lending banks.

2.2 Labor Supply and Wage Setting

The labor market structure is modeled following Schmitt-Grohe and Uribe (2006). In both countries, households of each type (patient, impatient) provide homogeneous labor services, which are transformed by monopolistically competitive unions into differentiated labor inputs. As a result, all households of the same type supply the same amount of hours worked in each sector, in equilibrium.

We assume that in each sector $j \in \{C, D\}$ there exist monopolistically competitive labor unions representing the patient and impatient households. Unions differentiate the homogeneous labor provided by households, N_{jt} from savers and $\widetilde{N}_{jt}$ from borrowers, creating a continuum of measure one of labor services (indexed by $z \in [0, 1]$) which are sold to labor packers.

[14]In our calibrations, the balanced growth path value of the loan-to-value ratio (LTV) $G(\overline{\varpi}_{HH,t})$ is around 0.5. This suggests that we would need shocks that cause extremely large movements in the LTV on impact before we violate the upper bound on the LTV. See the appendix in Bernanke, Gertler, and Gilchrist (1999) for a discussion of the same issue in their model.

[15]We use the non-residential goods price level as a deflator.

Then perfectly competitive labor packers buy the differentiated labor input and aggregate them through a CES technology into one labor input per sector and households type. Finally, the labor inputs are further combined using a Cobb-Douglas technology to produce the aggregate labor resource $L_{C,t}$ and $L_{D,t}$ that enter the production functions of entrepreneurs (see later). Unions set wages on a staggered basis. Every period, each union faces a constant probability $1 - \alpha_{wji}$ of being able to adjust its nominal wage. If the union is not allowed to reoptimize, wages are indexed to past and steady-state inflation. Taking into account that unions might not be able to choose their nominal wage optimally in the future, the optimal nominal wage is chosen to maximize intertemporal utility under the budget constraint and the labor demand function.

2.3 Non-Financial Corporate Sectors

2.3.1 Entrepreneurs

Entrepreneurs are also more impatient than household savers and have a discount factor $\beta_E < \beta$. They receive utility from their consumption of non-residential goods. They are in charge of the production of intermediate residential and non-residential goods, and operate in a perfectly competitive environment. They do not supply labor services. Their intertemporal utility function is given by

$$\mathcal{W}_t^E = \mathbb{E}_t \left\{ \sum_{j \geq 0} (\beta_E)^j \frac{\left(C_{t+j}^E - h_E C_{t+j-1}^E\right)^{1-\sigma_{CE}}}{1 - \sigma_{CE}} \varepsilon_{t+j}^{\beta} \right\}.$$

Non-residential intermediate goods are produced with capital and labor, while residential intermediate goods combine capital, labor, and land. In every period of time, savers are endowed with a given amount of land, which they sell to the entrepreneurs in a fixed quantity. We assume that the supply of land is exogenously fixed and that each entrepreneur takes the price of land as given in its decision problem. Entrepreneurs make use of Cobb-Douglas technology as follows:

$$Z_t(e) = \varepsilon_t^A \left(u_t^C(e) K_{t-1}^C(e)\right)^{\alpha_C} L_t^C(e)^{1-\alpha_C} - \Omega_C \quad \forall e \in [0,1]$$

$$Z_{D,t}(e) = \varepsilon_t^{A_D} \left(u_t^D(e) K_{t-1}^D(e)\right)^{\alpha_D} L_t^D(e)^{1-\alpha_D-\alpha_{\mathcal{L}}} \mathcal{L}_t(e)^{\alpha_{\mathcal{L}}} - \Omega_D,$$

where ε_t^A and $\varepsilon_t^{A_D}$ are exogenous technology shocks and $\mathcal{L}_t(e)$ denotes the endowment of land used by entrepreneur e at time t. Capital is sector specific and is augmented by a variable capacity utilization rate u_t. MC_t and $MC_{D,t}$ denote the selling prices for intermediate non-residential and residential products.

Entrepreneurs' fixed capital is subject to common multiplicative idiosyncratic shocks $\varpi_{E,t}$. As for households, these shocks are independent and identically distributed across time and across entrepreneurs with $E(\varpi_{E,t}) = 1$, and a log-normal CDF $F^E(\varpi_{E,t})$. Here again, the variance of the idiosyncratic shock $\sigma_{E,t}$ is time varying.

As for borrowers, entrepreneurs only use debt contracts in which the loan rates can be made contingent on aggregate shocks but not on the idiosyncratic shock $\varpi_{E,t}$. Entrepreneurs belong to a large family that can diversify the idiosyncratic risk after loan contracts are settled but cannot commit to sharing the proceeds of this insurance with banks. Banks can seize collateral $\varpi_{E,t}\tilde{A}_{E,t}$ when the entrepreneur refuses to pay at a cost of $\mu_E\varpi_{E,t}\tilde{A}_{E,t}$. The value of the collateral that the bank can seize is

$$\varpi_{E,t}\tilde{A}_{E,t} = \varpi_{E,t}(1-\chi_E)(1-\delta_K)\big(Q_t^C K_{t-1}^C + Q_t^D K_{t-1}^D\big).$$

We assume that the capital utilization rate is predetermined with respect to the idiosyncratic shock to facilitate aggregation. χ_E reflects the ability to collateralize capital. This specification relates to models where only capital serves as collateral, as in Kobayashi, Nakajima, and Inaba (2007) or Gerali et al. (2010).

Aggregate repayments or defaults on outstanding loans to entrepreneurs are

$$H^E(\overline{\varpi}_{E,t})\tilde{A}_{E,t} = \left[\big(1-F_t^E(\overline{\varpi}_{E,t})\big)\overline{\varpi}_{E,t} + \int_0^{\overline{\varpi}_{E,t}} \varpi dF_t^E\right]\tilde{A}_{E,t}.$$

On the commercial lending bank side, the profit made on the credit allocation is given by

$$G^E(\overline{\varpi}_{E,t})\tilde{A}_{E,t} - \frac{(1+R_{E,t-1})}{1+\pi_t}B_{E,t-1} \geq 0$$

with

$$G^E(\overline{\varpi}_{E,t}) = \big(1 - F_t^E(\overline{\varpi}_{E,t})\big)\overline{\varpi}_{E,t} + (1-\mu_E)\int_0^{\overline{\varpi}_{E,t}} \varpi dF_t^E.$$

$R_{E,t-1}$ is the interest rate at which the commercial lending bank gets financing every period, while $R_{E,t}^L$ is the state-contingent lending rate to entrepreneurs.

Overall, each entrepreneur maximizes its utility function with respect to $(C_t^E, K_t^C, K_t^D, u_t^C, u_t^D, B_t^E, \overline{\varpi}_{E,t}, L_{C,t}, L_{D,t})$ subject to an infinite sequence of real budget constraints

$$\begin{aligned} C_t^E &+ Q_t^C\big(K_t^C - (1-\delta_K)K_{t-1}^C\big) + Q_t^D\big(K_t^D - (1-\delta_K)K_{t-1}^D\big) \\ &\quad + H^E(\overline{\varpi}_{E,t})\tilde{A}_{E,t} \\ &= B_{E,t} + MC_tZ_t + MC_{D,t}Z_{D,t} - W_{C,t}^r L_{C,t} - W_{D,t}^r L_{D,t} - p_{lt}\mathcal{L}_t \\ &\quad - \Phi\big(u_t^C\big)K_{t-1}^C - \Phi\big(u_t^D\big)K_{t-1}^D + TT_t^E \end{aligned}$$

together with the participation constraints for the banks. We assume the following functional form for the adjustment costs on capacity utilization: $\Phi(X) = \frac{\overline{\overline{R^k}}(1-\varphi)}{\varphi}\left(\exp\left[\frac{\varphi}{1-\varphi}(X-1)\right] - 1\right)$. Following Smets and Wouters (2007), the cost of capacity utilization is zero when capacity is fully used ($\Phi(1) = 0$). p_{lt} denotes the relative price of land deflated by non-residential goods price.

2.3.2 Retailers and Distribution Sectors

Retailers differentiate the residential and non-residential goods produced by the entrepreneurs and operate under monopolistic competition. They sell their output to the perfectly competitive distribution sectors which aggregate the continuum of differentiated goods. The elementary differentiated goods are imperfect substitutes, with elasticity of substitution denoted $\frac{\mu_D}{\mu_D - 1}$ and $\frac{\mu}{\mu-1}$ for the residential and the non-residential sectors, respectively. The distributed goods are then produced with the following technology: $Y_D = \left[\int_0^1 Z_D(d)^{\frac{1}{\mu_D}} \mathrm{d}d\right]^{\mu_D}$ and $Y = \left[\int_0^1 Z(c)^{\frac{1}{\mu}} \mathrm{d}c\right]^{\mu}$. The corresponding aggregate price indexes are defined as $P_D = \left[\int_0^1 p_D(d)^{\frac{1}{1-\mu_D}} \mathrm{d}d\right]^{1-\mu_D}$

for the residential sector and $P = \left[\int_0^1 p(c)^{\frac{1}{1-\mu}} \mathrm{d}c\right]^{1-\mu}$ for the non-residential sector. The distribution goods serve as final consumption goods for households and are used by capital and housing stock producers.

Retailers are monopolistic competitors which buy the homogenous intermediate products of the entrepreneurs at prices MC_t for the non-residential intermediate goods and $MC_{D,t}$ for the residential intermediate goods. The intermediate products are then differentiated and sold back to the distributors. Retailers set their prices on a staggered basis à la Calvo (1983). In each period, a retailer in the non-residential sector faces a constant probability $1 - \xi_C$ (resp. $1 - \xi_D$ in the residential sector) of being able to reoptimize its nominal price. The demand curves that retailers face in each sector follow $Z_D(d) = \left(\frac{p_D(d)}{P_D}\right)^{-\frac{\mu_D}{\mu_D - 1}} Y_D$ and $Z(c) = \left(\frac{p(c)}{P}\right)^{-\frac{\mu}{\mu-1}} Y$.

2.3.3 Capital and Housing Stock Producers

Using distributed residential and non-residential goods, a segment of perfectly competitive firms, owned by the patient households, produce a stock of housing and fixed capital. At the beginning of period t, those firms buy back the depreciated housing stocks from both households types $(1 - \delta)D_{t-1}$ and $(1 - \delta)\widetilde{D}_{t-1}$ as well as the depreciated capital stocks $(1 - \delta_K)K^C_{t-1}$, $(1 - \delta_K)K^D_{t-1}$ at real prices (in terms of consumption goods) $Q_{D,t}T_{D,t}, \widetilde{Q}_{D,t}T_{D,t}, Q^D_t, Q^C_t$, respectively. Then they augment the various stocks using distributed goods and facing adjustment costs. The augmented stocks are sold back to entrepreneurs and households at the end of the period at the same prices.

2.4 The Banking Sector

The banking sector is owned by the patient households and is segmented in three parts. Following Gerali et al. (2010), each banking group is first composed of a wholesale branch which gets financing in the money market and allocates funds to the rest of the group, facing an adjustment cost on the overall capital ratio of the group. The wholesale branch takes the bank capital and the dividend policy as given in its decision problem and operates under perfect competition.

The second segment of the banking group comprises a deposit branch which collects savings from the patient households and places them in the money markets as well as two loan book financing branches which receive funding from the wholesale branch and allocate them to the commercial lending branches. In this second segment, banks operate under monopolistic competition and face nominal rigidity in their interest rate settings. The third segment of the banking group is formed by two commercial lending branches which provide loan contracts to impatient households and entrepreneurs. The commercial lending branches are zero-profit competitive firms.

2.4.1 Wholesale Branch

The perfectly competitive wholesale branches receive deposits Dep_t^{wb}, from the retail deposit banks, with an interest rate set at the policy rate R_t. Taking as given the bank capital $Bankcap_t$ in real terms, they provide loans $B_{E,t}^{wb}$ and $B_{HH,t}^{wb}$ at interest rates $R_{E,t}^{wb}$ and $R_{HH,t}^{wb}$ to the loan book financing branches for lending to entrepreneurs and households, respectively. When deciding on deposits and loans, the wholesale banks are constrained by an adjustment cost on banks' leverage. This friction is meant to capture the capital requirement pressures on the banks' behavior. For this reason, we assume that wholesale banks target a capital ratio of 11 percent and the quadratic cost is supposed to illustrate the various interactions between banks' balance sheet structure, market disciplining forces, and the regulatory framework.[16] On the one hand, this reflects that, owing to pecuniary and reputational costs, banks are keen to avoid getting too close to the regulatory minimum capital requirement and hence tend to operate with a substantial buffer over that minimum capital ratio.[17] On the other hand, bank capital is costly relative to other sources of financing (like deposits and bond issuance),

[16]The 11 percent capital ratio target corresponds to the average (risk-adjusted) total capital ratio of the 100 largest listed euro-area banks for the period 1999–2008 according to Datastream (Worldscope).

[17]There is a rich literature providing evidence that banks operate with substantial capital buffers; for some recent studies see, e.g., Ayuso, Pérez, and Saurina (2004), Bikker and Metzemakers (2004), Berger et al. (2008), Stolz and Wedow (2005), and Gropp and Heider (2010).

implying that banks tend to economize on the amount of capital they hold.[18]

Under the Basel I-like capital requirement regime, the bank's static profit maximization problem can be formulated as follows, where all quantities are expressed in real terms:

$$\max_{B_t^w, Dep_t^w} R_{HH,t}^{wb} B_{HH,t}^{wb} + R_{E,t}^{wb} B_{E,t}^{wb} - R_t Dep_t^{wb} - \frac{\chi_{wb}}{2} \left(\frac{Bankcap_t}{0.5 B_{HH,t}^{wb} + B_{E,t}^{wb}} - 0.11 \right)^2 Bankcap_t$$

subject to the balance sheet identity

$$B_{HH,t}^{wb} + B_{E,t}^{wb} = Dep_t^{wb} + Bankcap_t.$$

As in Gerali et al. (2010) the derived lending spreads emphasize "the role of bank capital in determining loan supply conditions." Hence, on the one hand, if the spread between the lending rate and the policy rate is positive, the bank would have an incentive to increase profits by raising loan volumes. This, on the other hand, would increase its leverage, which is, however, penalized by regulatory rules and market disciplining forces, as the capital ratio moves away from its target, which poses a cost to the bank. The bank's decision problem is therefore finely balanced between boosting its profits via increased leverage and retaining control of its capital structure. Moreover, a key point to notice for our Basel I type specification is that the bank's target capital ratio is insensitive to changes in borrower risk over time. In addition, reflecting the risk weighting of the Basel I regulatory framework, household loans are given a (fixed) risk weight of 50 percent, whereas the risk weight attached to corporate loans is 100 percent.

[18]For example, the European Central Bank (ECB) estimates of the cost of equity, the cost of market-based debt (i.e., bond issuance), and the cost of deposits for euro-area banks show that the former was on average around 6.7 percent in the period 2003–09. During the same period, banks' cost of raising debt in the capital markets was around 5 percent, while their average cost of deposit funding was close to 2 percent.

The decision problem of the wholesale bank leads to the following condition on the spread between the lending rate and the policy rate:

$$R^{wb}_{HH,t} - R_t = -\chi_{wb}\left(\frac{Bankcap_t}{0.5B^{wb}_{HH,t} + B^{wb}_{E,t}} - 0.11\right) \times \left(\frac{Bankcap_t}{0.5B^{wb}_{HH,t} + B^{wb}_{E,t}}\right)^2 0.5$$

$$R^{wb}_{E,t} - R_t = -\chi_{wb}\left(\frac{Bankcap_t}{0.5B^{wb}_{HH,t} + B^{wb}_{E,t}} - 0.11\right) \times \left(\frac{Bankcap_t}{0.5B^{wb}_{HH,t} + B^{wb}_{E,t}}\right)^2.$$

When the leverage of the bank increases beyond the targeted level, banks increase their loan-deposit margins.

The capital base of the wholesale branch is accumulated out of retained earnings from the bank group profits

$$Bankcap_t = (1 - \delta^{wb})Bankcap_{t-1} + \nu^b \Pi^b_t,$$

where δ^{wb} represents the resources used in managing bank capital, Π^b_t is the overall profit of the bank group, and ν^b is the share of profits not distributed to the patient households.

2.4.2 Imperfect Pass-Through of Policy Rate on Bank Lending Rates

The retail deposit branch and the loan book financing branches are monopolistic competitors and set their interest rates on a staggered basis with some degree of nominal rigidity à la Calvo (1983).

Retail Deposit Branch. The deposits offered to patient households are a CES aggregation of the differentiated deposits provided by the retail deposit branches: $Dep = \left[\int_0^1 Dep(j)^{\frac{1}{\mu^R_D}} \mathrm{d}j\right]^{\mu^R_D}$, expressed in real terms. Retail deposits are imperfect substitutes with elasticity of substitution $\frac{\mu_D^R}{\mu_D^R - 1} < -1$. The corresponding average interest rate offered on deposits is $R_D = \left[\int_0^1 R_D(j)^{\frac{1}{1-\mu_D^R}} \mathrm{d}j\right]^{1-\mu_D^R}$.

Retail deposit branches are monopolistic competitors which collect deposits from savers and place them in the money market. Deposit branches set interest rates on a staggered basis à la Calvo (1983), facing each period a constant probability $1 - \xi_D^R$ of being able to reoptimize their nominal interest rate. When a retail deposit branch cannot reoptimize its interest rate, the interest rate is left at its previous period level.

The retail deposit branch j chooses $\hat{R}_{D,t}(j)$ to maximize its intertemporal profit.

$$\mathbb{E}_t \left[\sum_{k=0}^{\infty} \left(\gamma \xi_D^R\right)^k \frac{\Lambda_{t+k}}{\Lambda_t} \left(R_{t+k} Dep_{t+k}(j) - \hat{R}_{t,D}(j) Dep_{t+k}(j)\right)\right],$$

where $Dep_{t+k}(j) = \left(\frac{\hat{R}_{D,t}(j)}{R_{D,t}}\right)^{-\frac{\mu_D^R}{\mu_D^R - 1}} \left(\frac{R_{D,t}}{R_{D,t+k}}\right)^{-\frac{\mu_D^R}{\mu_D^R - 1}} Dep_{t+k}$ and Λ_t is the marginal value of non-residential consumption for the household savers.

A markup shock $\varepsilon_{D,t}^R$ is introduced in the staggered nominal deposit rate setting.

Loan Book Financing Branches. As for the retail deposit branches, loan book financing branches provide funds to the commercial lending branches which obtain overall financing through a CES aggregation of the differentiated loans: $B_{E,t} = \left[\int_0^1 B_{E,t}(j)^{\frac{1}{\mu_E^R}} \mathrm{d}j\right]^{\mu_E^R}$ as regards commercial loans to entrepreneurs and $B_{HH,t} = \left[\int_0^1 B_{HH,t}(j)^{\frac{1}{\mu_{HH}^R}} \mathrm{d}j\right]^{\mu_{HH}^R}$ as regards commercial loans to households. Loans from loan book financing branches are imperfect substitutes with elasticity of substitution $\frac{\mu_E^R}{\mu_E^R - 1}$ and $\frac{\mu_{HH}^R}{\mu_{HH}^R - 1} > 1$. The corresponding average lending rate is

$$R_E = \left[\int_0^1 R_E(j)^{\frac{1}{1-\mu_E^R}} \mathrm{d}j\right]^{1-\mu_E^R} \quad \text{and}$$

$$R_{HH} = \left[\int_0^1 R_{HH}(j)^{\frac{1}{1-\mu_{HH}^R}} \mathrm{d}j\right]^{1-\mu_{HH}^R}.$$

Loan book financing branches for each segment of the credit market are monopolistic competitors which levy funds from the wholesale branches and set interest rates on a staggered basis à la Calvo (1983), facing each period a constant probability $1-\xi_E^R$ and $1-\xi_{HH}^R$ of being able to reoptimize their nominal interest rate. If a loan book financing branch cannot reoptimize its interest rate, the interest rate is left at its previous period level.

In each sector $i \in \{E, HH\}$, the loan book financing branch j chooses $\hat{R}_{i,t}(j)$ to maximize its intertemporal profit.

$$\mathbb{E}_t \left[\sum_{k=0}^{\infty} \left(\gamma \xi_i^R\right)^k \frac{\Lambda_{t+k}}{\Lambda_t} \left(\hat{R}_{i,t}(j) B_{i,t+k}(j) - R_{i,t}^{wb}(j) B_{i,t+k}(j)\right) \right],$$

where $B_{i,t+k}(j) = \left(\frac{\hat{R}_{i,t}(j)}{R_{i,t}}\right)^{-\frac{\mu_i^R}{\mu_i^R-1}} \left(\frac{R_{i,t}}{R_{i,t+k}}\right)^{-\frac{\mu_i^R}{\mu_i^R-1}} B_{i,t+k}$.

As for deposit rates, we add markup shocks $\varepsilon_{HH,t}^R$ and $\varepsilon_{E,t}^R$ to the staggered nominal lending rate settings.

Commercial Lending Branches. Commercial lending branches deliver credit contracts for entrepreneurs and household borrowers. Those branches are perfectly competitive and in equilibrium have zero profits. Details on the credit contract and the decision problems for the commercial lending branches are provided in the sections on entrepreneurs and household borrowers.

2.5 Government and Monetary Authority

Public expenditures $\overline{\overline{G}}$ are subject to random shocks ε_t^G. The government finances public spending with lump-sum transfers.

Monetary policy is specified in terms of an interest rate rule targeting inflation, output, and their first difference as well as changes in the relative price of housing. Written in deviation from the steady state, the interest rate rule used has the following form:

$$\begin{aligned} r_t = {} & \rho r_{t-1} + (1-\rho)(r_\pi \pi_{t-1} + r_y y_{t-1}) + r_{\Delta\pi} \Delta\pi_t + r_{\Delta y} \Delta y_t \\ & + r_{T_D} \Delta t_{D,t} + \log\left(\varepsilon_t^R\right), \end{aligned}$$

where lowercase letters denote log-deviations of a variable from its deterministic steady state.

3. Bayesian Estimation

The model is estimated on euro-area data using Bayesian likelihood methods. We consider fifteen key macroeconomic quarterly time series from 1986:Q1 to 2008:Q2: output, consumption, non-residential fixed investment, hours worked, real wages, CPI inflation rate, three-month short-term interest rate, residential investment, real house prices, household loans, non-financial corporation loans, household deposits, and bank lending rates on household loans, on non-financial corporation loans, and on household deposits. All real variables and real house prices are linearly detrended prior to estimation. Inflation and nominal interest rates are mean adjusted (see the calibration section for more details). Full description of the data set is provided in appendix 1.

We summarize here the exogenous stochastic shocks that we introduce:

- Efficient shocks: AR(1) technology (ε_t^A) (common to both sectors), AR(1) housing-specific technology ($\varepsilon_t^{A_D}$), AR(1) non-residential investment specific productivity (ε_t^I), AR(1) labor supply (ε_t^L), AR(1) public expenditure (ε_t^G), AR(1) consumption preferences (ε_t^B), and AR(1) housing preferences (ε_t^D)
- Inefficient shocks: i.i.d price markup (ε_t^P), AR(1) interest rate markups on deposits and loans ($\varepsilon_{D,t}^R, \varepsilon_{HH,t}^R, \varepsilon_{E,t}^R$).
- Riskiness shocks: the standard deviation of the idiosyncratic risk for impatient households and entrepreneurs is subject to AR(1) shocks ($\varepsilon_{HH,t}^\sigma, \varepsilon_{E,t}^\sigma$)
- AR(1) bank capital shock ($\varepsilon_t^{Bankcap}$)
- Monetary policy shock (ε_t^R)

As regards behavioral parameters, we chose to limit the number of estimated coefficients by bringing some symmetry across sectors and agents. We estimate the parameters driving the adjustment costs on residential and non-residential investment, ϕ_D, ϕ, which are the same across household types and sectors, respectively. The parameter on capacity utilization adjustment cost φ is also the same for both sectors. Concerning preference parameters, the intertemporal elasticity of substitution, σ_X, is similar for the two household types; the labor supply elasticity, σ_L, is the same across household types and sector-specific labor service; and the habit parameter, h, is equalized across all agents. The Calvo parameters on nominal wage rigidity, α_{wC}, α_{wD}, are the same for both household types, while we introduce a single indexation parameter γ_w. The Calvo parameter on non-residential retail goods price setting, ξ_C, and the associated indexation coefficients, γ_C, are estimated, while in the residential goods sector, we estimate the Calvo parameter, ξ_D, and set the indexation parameter, γ_D, to zero. On the imperfect interest rate pass-through, we draw some inference on the three coefficients driving the staggered rate setting on deposits and loans, ξ_D^R, ξ_{HH}^R, ξ_E^R. The adjustment cost on banks' capital structure, χ_{wb}, is also estimated. Finally, the parameters in the Taylor rule are $\rho, r_\pi, r_y, r_{\Delta\pi}, r_{\Delta y}, r_{T_D}$.

In the benchmark estimation, we do not introduce the share of household borrowers. As argued later on, given the weak identification of the parameter and the lack of observable data on households' heterogenous features, we calibrated this parameter to achieve realistic debt structure in the steady state. At the same, some inference and sensitivity analysis on this coefficient is presented thereafter. Calibrating the share of borrowers is also symmetric to our assumption that all firms are financially constrained.

Some parameters are excluded from the estimation and have to be calibrated. These are typically parameters driving the steady-state values of the state variables, for which the econometric model based on detrended data is almost non-informative. Details about the calibrated parameters, the steady state, and the prior distributions are provided in appendices 2 and 3.

3.1 Posterior Distributions

Tables 1 and 2 report the mode, the mean, and the 10th and 90th percentiles of the posterior distribution of the structural parameters for the model.

In terms of the parameter estimates, emphasizing those features that are more closely related to our modeling framework with respect to the sectoral structure of the economy and financial frictions, among the stochastic exogenous disturbances, the posterior distributions for autoregressive coefficients turned out to be very close to unity for several shocks. Those shocks are, notably, those related to the housing sector, housing preference and productivity shocks,

Table 1. Parameter Estimates 1

Param.	*A priori* Beliefs			Benchmark Specification *A posteriori* Beliefs				Alternative Specification *A posteriori* Beliefs			
	Dist.	Mean	Std.	Mode	Mean	$\mathcal{I}_1$	$\mathcal{I}_2$	Mode	Mean	$\mathcal{I}_1$	$\mathcal{I}_2$
ε_t^A	unif	5	2.89	0.49	0.50	0.42	0.58	0.45	0.46	0.39	0.53
ε_t^I	unif	5	2.89	0.26	0.27	0.20	0.34	0.26	0.28	0.20	0.35
ε_t^L	unif	5	2.89	0.17	0.18	0.13	0.23	0.18	0.19	0.15	0.24
ε_t^G	unif	5	2.89	1.85	1.87	1.64	2.12	1.88	1.92	1.68	2.16
ε_t^B	unif	5	2.89	1.44	1.50	1.24	1.77	2.05	2.02	1.57	2.45
$\varepsilon_t^{A_D}$	unif	5	2.89	2.09	2.15	1.52	2.75	2.08	2.10	1.50	2.65
ε_t^D	unif	5	2.89	2.04	2.47	1.21	3.72	2.77	2.95	1.72	4.10
ε_t^P	unif	5	2.89	0.28	0.30	0.25	0.35	0.28	0.29	0.25	0.34
$\varepsilon_{D,t}^R$	unif	5	2.89	0.06	0.06	0.04	0.08	0.07	0.07	0.05	0.09
$\varepsilon_{HH,t}^R$	unif	5	2.89	0.06	0.06	0.05	0.07	0.06	0.06	0.05	0.07
$\varepsilon_{E,t}^R$	unif	5	2.89	0.12	0.13	0.11	0.15	0.12	0.13	0.11	0.15
$\varepsilon_{HH,t}^\sigma$	unif	5	2.89	0.08	0.08	0.07	0.09	0.03	0.03	0.02	0.04
$\varepsilon_{E,t}^\sigma$	unif	5	2.89	0.06	0.06	0.05	0.07	0.06	0.06	0.05	0.07
$\varepsilon_t^{Bankcap}$	unif	5	2.89	2.32	2.37	2.06	2.66	2.30	2.35	2.06	2.65
ε_t^R	unif	5	2.89	0.10	0.11	0.09	0.12	0.11	0.11	0.10	0.13

(continued)

Table 1. (Continued)

Param.	*A priori* Beliefs			Benchmark Specification *A posteriori* Beliefs				Alternative Specification *A posteriori* Beliefs			
	Dist.	**Mean**	**Std.**	**Mode**	**Mean**	$\mathcal{I}_1$	$\mathcal{I}_2$	**Mode**	**Mean**	$\mathcal{I}_1$	$\mathcal{I}_2$
ρ_A	beta	0.5	0.2	0.92	0.90	0.85	0.95	0.97	0.96	0.93	0.98
ρ_I	beta	0.5	0.2	0.68	0.68	0.58	0.78	0.69	0.68	0.58	0.79
ρ_l	beta	0.5	0.2	0.85	0.83	0.73	0.93	0.79	0.60	0.27	0.89
ρ_G	beta	0.5	0.2	0.91	0.88	0.79	0.97	0.99	0.98	0.97	0.99
ρ_B	beta	0.5	0.2	0.95	0.94	0.91	0.97	0.97	0.97	0.96	0.98
ρ_{A_D}	beta	0.5	0.2	0.89	0.88	0.83	0.94	0.90	0.89	0.84	0.93
ρ_D	beta	0.5	0.2	0.99	0.98	0.97	0.99	0.97	0.97	0.96	0.99
$\rho^R_{D,t}$	beta	0.5	0.2	0.97	0.96	0.93	0.99	0.95	0.95	0.92	0.97
$\rho^R_{HH,t}$	beta	0.5	0.2	0.26	0.26	0.13	0.38	0.25	0.25	0.14	0.37
$\rho^R_{E,t}$	beta	0.5	0.2	0.40	0.41	0.23	0.57	0.37	0.37	0.22	0.52
$\rho^\sigma_{HH,t}$	beta	0.5	0.2	0.98	0.98	0.97	0.99	0.99	0.99	0.98	0.99
$\rho^\sigma_{E,t}$	beta	0.5	0.2	0.98	0.98	0.96	0.99	0.98	0.97	0.96	0.99
$\rho_t^{Bankcap}$	beta	0.5	0.2	0.71	0.70	0.61	0.79	0.72	0.71	0.62	0.80

Table 2. Parameter Estimates 2

Param.	*A priori* Beliefs			Benchmark Specification				Alternative Specification			
				A posteriori Beliefs				*A posteriori* Beliefs			
	Dist.	Mean	Std.	Mode	Mean	$\mathcal{I}_1$	$\mathcal{I}_2$	Mode	Mean	$\mathcal{I}_1$	$\mathcal{I}_2$
ϕ_D	gamm	1	0.5	0.23	0.23	0.15	0.31	0.23	0.22	0.15	0.30
ϕ	norm	4	1.5	7.72	7.71	5.63	9.74	7.52	7.66	5.54	9.79
φ	beta	0.5	0.15	0.75	0.76	0.63	0.89	0.77	0.76	0.64	0.89
σ_X	gamm	1.5	0.20	0.63	0.65	0.56	0.73	0.71	0.72	0.66	0.79
h	beta	0.75	0.1	0.59	0.58	0.54	0.62	0.53	0.51	0.45	0.58
σ_L	gamm	1.5	0.1	1.32	1.33	1.18	1.47	1.32	1.32	1.18	1.46
α_{wC}	beta	0.85	0.05	0.73	0.73	0.67	0.80	0.73	0.79	0.68	0.89
α_{wD}	beta	0.85	0.05	0.87	0.85	0.79	0.92	0.88	0.86	0.81	0.92
γ_w	beta	0.5	0.15	0.24	0.28	0.10	0.44	0.19	0.26	0.09	0.43
ξ_C	beta	0.75	0.05	0.83	0.83	0.79	0.86	0.84	0.85	0.81	0.88
γ_C	beta	0.5	0.15	0.58	0.61	0.45	0.78	0.55	0.58	0.43	0.73
ξ_D	beta	0.2	0.1	0.81	0.81	0.76	0.86	0.81	0.79	0.75	0.84
ξ_D^R	beta	0.5	0.2	0.32	0.31	0.26	0.37	0.28	0.28	0.23	0.33

(continued)

Table 2. (Continued)

Param.	*A priori* Beliefs			Benchmark Specification *A posteriori* Beliefs				Alternative Specification *A posteriori* Beliefs			
	Dist.	Mean	Std.	Mode	Mean	$\mathcal{I}_1$	$\mathcal{I}_2$	Mode	Mean	$\mathcal{I}_1$	$\mathcal{I}_2$
ξ^R_{HH}	beta	0.5	0.2	0.92	0.92	0.88	0.95	0.90	0.90	0.88	0.92
ξ^R_E	beta	0.5	0.2	0.76	0.75	0.69	0.82	0.77	0.76	0.70	0.82
χ_{wb}	gamm	20	2.5	18.58	18.54	15.08	22.03	19.07	19.34	15.90	22.56
ρ	beta	0.75	0.1	0.84	0.84	0.82	0.87	0.81	0.81	0.78	0.84
r_π	gamm	2.5	0.25	2.37	2.38	2.17	2.59	2.36	2.41	2.22	2.60
r_y	gamm	0.2	0.1	0.03	0.03	0.01	0.05	0.04	0.04	0.02	0.06
$r_{\Delta\pi}$	gamm	0.30	0.10	0.24	0.24	0.18	0.31	0.31	0.31	0.24	0.38
$r_{\Delta y}$	gamm	0.12	0.05	0.07	0.08	0.05	0.11	0.09	0.10	0.06	0.13
r_{T_D}	norm	0.00	1.00	0.03	0.03	0.00	0.06	0.01	0.01	−0.02	0.04
λ_e	beta	0.75	0.05	0.65	0.65	0.61	0.69	0.64	0.63	0.58	0.67
ω	beta	0.45	0.05	—	—	—	—	0.28	0.27	0.22	0.31
$\rho_{B,D}$	unif	0.00	2.89	—	—	—	—	0.89	0.91	0.53	1.26
$\rho_{D,\sigma_{HH}}$	gamm	1.00	0.50	—	—	—	—	1.78	1.76	1.20	2.33
$P_\lambda(\mathcal{Y})$				−432.1				−387.2			

and to loan dynamics, risk shocks on households and entrepreneurs. Visual inspection of detrended real house prices and loan data over the sample indeed suggest very high degrees of persistence which are not well captured by the internal propagation of the model. The markup shocks on bank interest rates also display high autoregressive coefficients, with the notable exception of lending rates to households, for which lower inertia seems to compensate for higher nominal rigidity.

Turning to behavioral parameters, the labor supply elasticity as well as the inflation term in level in the monetary policy rule are weakly identified. The estimation does not support the evidence of meaningful specific reaction of monetary policy to house prices.

The Calvo parameters on the imperfect adjustment of lending rates are estimated to be the lowest for deposit rates, at around 0.3; the highest for lending rates to households, at around 0.9 in the benchmark estimation; and somewhat in between for lending rates to entrepreneurs, at around 0.75. The higher flexibility of deposit rates is also found by Gerali et al. (2010) and is most likely due to differences in the maturity structures of the various composite rates which cannot be accounted for by the one-period loans considered in the DSGE model.

Finally, the posterior distribution for the adjustment cost on banks' capital structure, χ_{wb}, stays very close to its prior distribution. At the same time, having experimented with alternative priors, the posterior distribution could eventually depart significantly from the prior one, therefore suggesting that data are somewhat informative about this parameter.

As regards the real and nominal rigidities for the residential sector, the estimation leads to an adjustment cost parameter for residential investment, ϕ_D, of around 0.2 at the mode. The degree of nominal rigidity is quite elevated, with a posterior mode for the Calvo parameter on residential prices of 0.81. The real rigidities in the residential sector have compounded effects on macroeconomic propagation through households' borrowing constraint and, consequently, households' consumption expenditures. Overall, it seems that data call for some degree of real rigidity in the residential markets. Everything else being equal, this implies that relative prices would react more to economic shocks. In order to limit the volatility of residential prices in the presence of adjustment costs on residential investment, staggered housing price setting is needed.

Tables 1 and 2 show the posterior parameter distributions when introducing correlations between the consumption preference shock and the housing preference shock on the one hand, and between the housing preference shocks and the household risk shock on the other hand. These experiments were guided by the correlations of structural shocks obtained in the estimations.

The innovation on the consumption preference shock, ε_t^B, has been introduced in the AR(1) process of the housing preference shock. Such a positive correlation between both exogenous disturbances is partly correcting for the sharp negative co-movement after a consumption preference shock between consumption and residential investment, which may not be supported by data given the positive unconditional correlation observed in our sample. The introduction of the innovation on the housing preference shock in the AR(1) process of the risk shock on housing loans is limiting the negative co-movement between residential price and residential investment on the one hand and lending rate spreads to households on the other hand. The presence of such correlations is affecting the inference on behavioral parameters.

In the results reported in tables 1 and 2, we also estimated the share of household borrowers, ω, free in the estimation procedure. The prior distribution for this parameter was set with a relatively elevated mean and small variance. The posterior distribution for the household borrowers' share reaches 28 percent at the mode. Overall, ω does not seem to be strongly identified. This confirms the results of Darracq Pariès and Notarpietro (2008). The presence of borrowers is not rejected by the data, as the model specification leads to strictly positive values for such shares.

3.2 Business-Cycle Contribution of Financial Shocks

We also analyze the role of credit market frictions and financial shocks in economic fluctuations. Table 3 reports unconditional variance decomposition of HP-filtered variables, emphasizing the contribution of housing-related structural shocks and shocks to the banking sector. The variance decomposition is computed using the posterior modes of their respective estimation.

Table 3. Shocks Decomposition of Unconditional Variances: HP Filtering

	$\varepsilon_t^{A_D}$	ε_t^{D}	$\varepsilon_{HH,t}^{\sigma}$	$\varepsilon_{E,t}^{\sigma}$	$\varepsilon_{HH,t}^{R}$	$\varepsilon_{E,t}^{R}$	$\varepsilon_{D,t}^{R}$	$\varepsilon_t^{Bankcap}$	**Others** Z_t
Z_t	3.1	19.1	14.9	3.5	1.2	0.1	8.2	0.7	49.1
C_t^{tot}	0.7	7.8	26.8	1.8	1.8	0.1	8.0	0.6	52.6
I_t	0.1	0.8	0.4	40.4	0.4	1.9	0.4	6.3	49.1
$Z_{D,t}$	23.3	42.9	1.4	0.8	0.4	0.1	5.0	0.1	26.1
$T_{D,t}$	3.6	36.5	1.3	3.0	0.0	0.1	1.2	0.5	53.9
L_t^{tot}	2.5	17.6	13.2	3.7	1.1	0.1	7.2	0.8	53.9
W_t^{tot}	0.5	4.3	16.0	1.2	1.1	0.0	6.9	0.4	69.7
Π_t	0.2	1.2	7.6	3.6	0.4	0.1	8.3	0.8	77.9
R_t	0.1	10.9	12.6	8.5	0.9	0.2	13.4	1.8	51.7
$R_{E,t}$	0.0	3.0	4.4	19.6	0.6	35.8	6.6	7.1	23.0
$R_{HH,t}$	0.2	8.5	24.6	2.7	42.9	0.8	5.9	2.6	11.9
$R_{D,t}$	0.1	11.6	14.4	10.5	0.9	0.3	4.7	2.2	55.4
$B_{E,t}$	0.0	1.5	6.3	61.1	0.8	1.0	3.5	3.1	22.7
$B_{HH,t}$	0.3	17.9	53.9	0.9	3.9	0.3	1.1	3.2	18.7
Dep_t	0.1	6.9	18.4	22.4	2.2	1.1	2.0	23.3	23.6

More than 50 percent of unconditional variances of loans to households and entrepreneurs are explained by their respective risk shock. Indeed, looking at zero-profit condition for household loans, for example,

$$G(\varpi_{HH,t})\tilde{A}_{HH,t} - \frac{(1 + R_{HH,t-1})}{1 + \pi_t} B_{HH,t-1} \geq 0,$$

we see that the term $G(\varpi_{HH,t})$ could be interpreted as a time-varying loan-to-value ratio and is directly related to the risk shock on household borrowers. In the empirical exercise, this shock is therefore partly capturing the gap between the dynamics of loans and the dynamics of its collateral value. Household deposits are mainly driven by risk shocks on households and entrepreneurs as well as by bank capital shocks, with a respective contribution of around 20 percent. Those disturbances have a strong impact on bank assets and capital, thereby mechanically affecting bank liabilities. Overall, approximately 20 percent of the unconditional volatility of loans and deposits are driven by disturbances not related to the financial or housing blocks.

On bank lending rates, for each sector, the risk shock and the interest rate markup shock have strong contributions, explaining jointly more than 50 percent of variance. By contrast, the role of financial shocks is more limited as regards the volatility of deposit rates.

Turning to the residential sector, the housing preference shock explains a large part of price and quantity in this sector. The housing-specific productivity shock contributes mainly to residential investment volatility. On balance, 40 percent of residential investment and 60 percent of real housing prices are driven by non-housing-specific disturbances.

For the non-residential sector, the corporate risk shock has a large contribution to non-residential investment fluctuations, whereas the household risk shock contributes significantly to consumption volatility, albeit to a lesser extent. The housing preference shock and the interest rate markup shock on deposits are non-negligible sources of consumption unconditional variance. For GDP, consumption, and non-residential investment, roughly 50 percent of unconditional variances are not explained by financial and housing-specific shocks.

Finally, on consumer prices, the risk shocks and the interest rate shock on deposits have some meaningful contributions, but almost 80 percent of variance is driven by disturbances not related to the financial or the housing blocks.

4. Macroeconomic Propagation and Monetary Policy Stabilization under Different Regulatory Frameworks

In this section we consider the macroeconomic implications of the various types of credit frictions embedded in our model and, in turn, consider how different kinds of regulatory frameworks might affect monetary policy stabilization.

We focus in particular on disturbances to the financial side of our model economy. For a description of propagation of non-financial economic disturbances, please refer to the working paper version of this article (Darracq Pariès, Kok Sørensen, and Rodriguez Palenzuela 2010).

4.1 Bank Capital Shocks and Bank Capital Channel

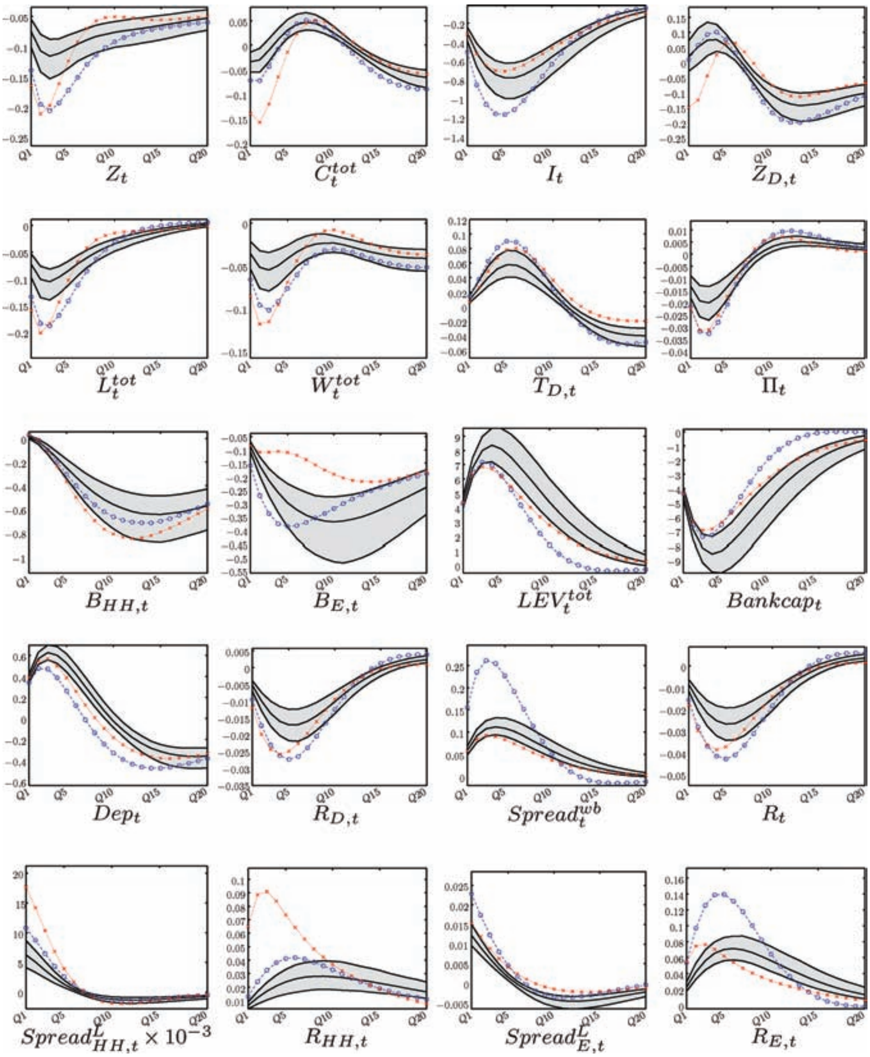
The recent financial crisis led banks to incur substantial losses on their trading and loan books, which in turn put severe pressure on their capital positions. In order to return to a more stable capital situation, and possibly responding to pressures from regulators and market participants to operate with more solid capital buffers, banks have been faced with a trade-off of either raising new capital or adjusting their asset side, or (more likely) a combination of the two. Our model specification can be used to assess the macroeconomic implications of such shocks to bank capital, which in our case will lead banks to replenish their capital position by boosting their retained earnings. This is illustrated in figure 2, which shows the implications of an adverse shock to bank capital, ($\varepsilon_t^{Bankcap}$). The bank capital shock results in an increase in bank leverage which, in order for banks to reestablish their target leverage ratio, leads to an increase in banks' loan-deposit margins. This is driven mainly by higher lending rates, which in turn reduces lending and hence real activity.[19] The negative impact on output of the bank capital shock in the benchmark model is relatively modest but persistent.[20]

The specific role of the bank capital channel in the propagation of economic shocks via the financial sector can be further analyzed by increasing banks' adjustment cost on their leverage (setting $\chi_{wb} = 50$). This is illustrated by the dotted lines with circles in figures 2–4, and it is observed that a more pronounced bank capital channel results in a much stronger propagation of shocks from the banking sector to the real economy. Consequently, the monetary policy response is also more forceful than in the benchmark case, which allows for output to rebound back towards the baseline over time.

[19] This mechanism is corroborated by empirical findings for the United States, which suggests that pressure on bank capital positions induces banks to apply higher lending rates (in particular vis-à-vis their riskier borrowers); see Santos and Winton (2009).

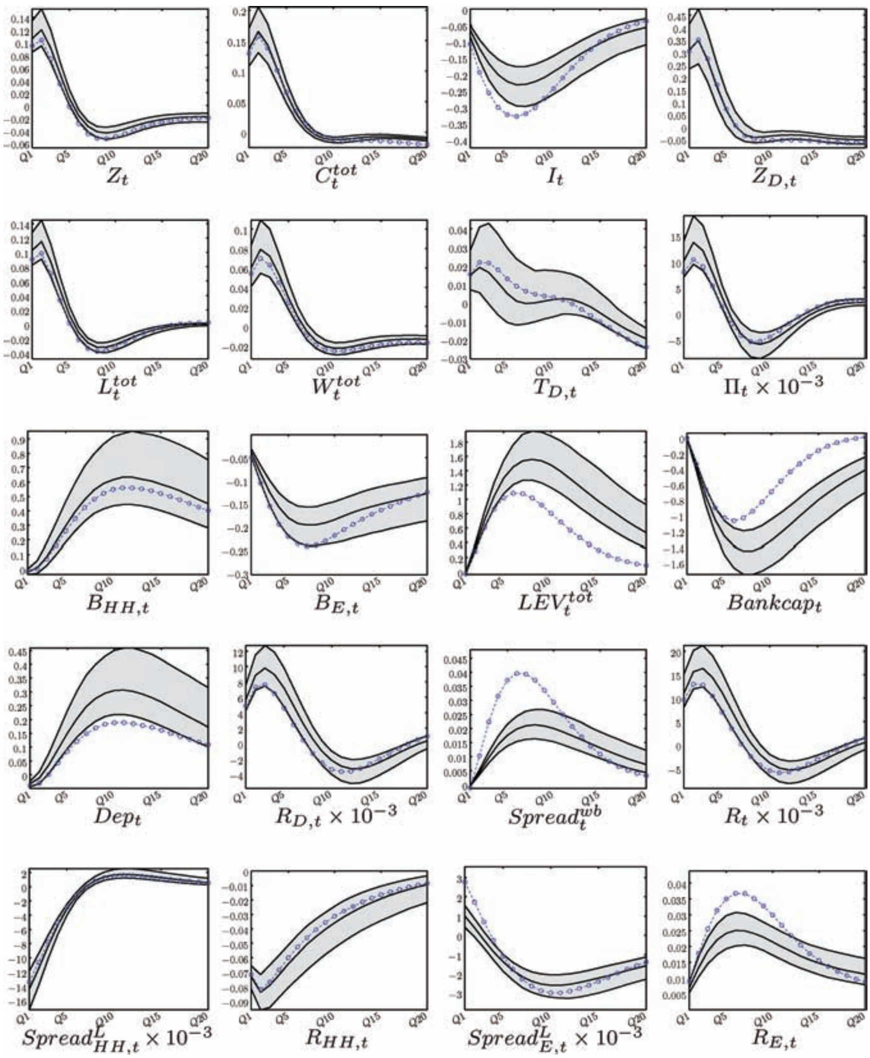
[20] Recent empirical studies suggest an approximate effect of a 1-percentage-point shock to bank capital positions (or loan supply shocks more generally) in the range of an approximately 0.1- to 1.0-percentage-point impact on real economic activity; see, e.g., Van den Heuvel (2008), Ciccarelli, Maddaloni, and Peydró (2009), Francis and Osborne (2009), and Cappiello et al. (2010). Our baseline estimates are at the lower end of this range.

Figure 2. Impulse Response Functions Associated with a Shock on $\varepsilon_t^{Bankcap}$



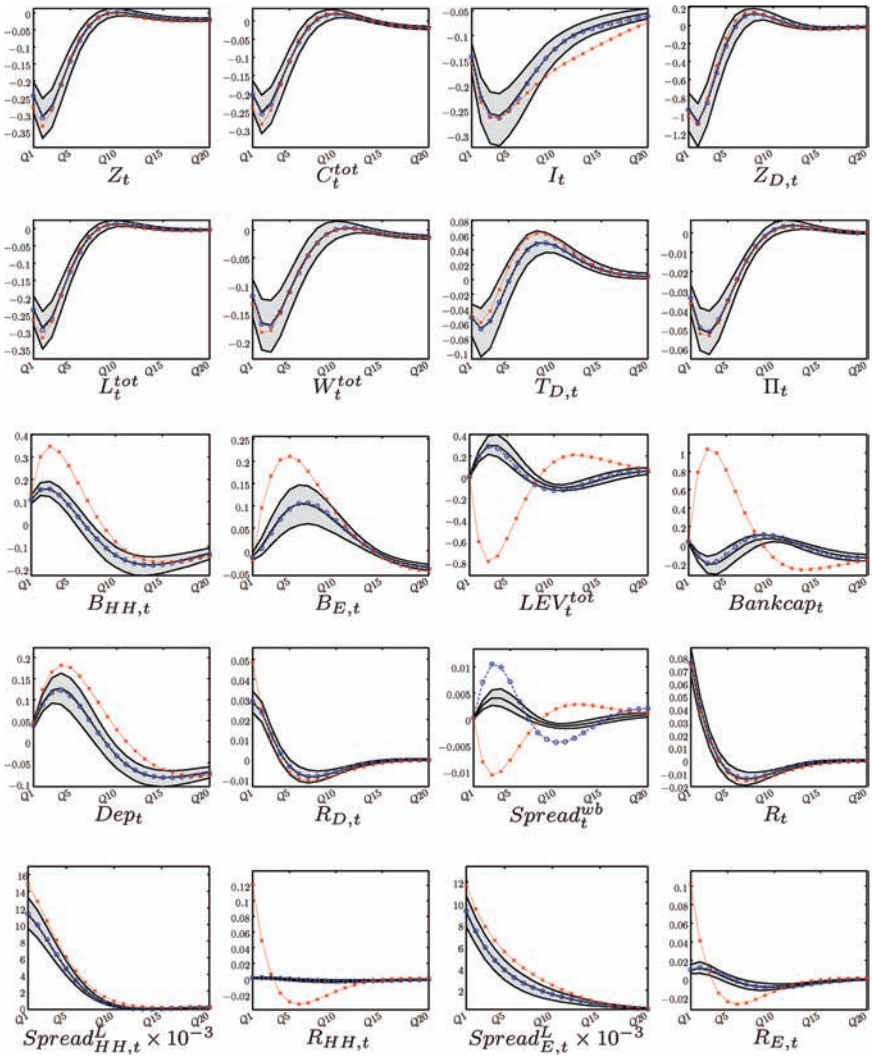
Note: Benchmark model (plain lines for the mode and shaded areas for the 90 percent highest-density interval from the posterior distribution of parameters), model with high bank capital channel (dotted lines with circles), and model without imperfect interest rate pass-through (dashed lines with crosses).

Figure 3. Impulse Response Functions Associated with a Shock on $\varepsilon^{R}_{HH,t}$



Note: Benchmark model (plain lines for the mode and shaded areas for the 90 percent highest-density interval from the posterior distribution of parameters), model with high bank capital channel (dotted lines with circles), and model without imperfect interest rate pass-through (dashed lines with crosses).

Figure 4. Impulse Response Functions Associated with a Shock on ε_t^R



Note: Benchmark model (plain lines for the mode and shaded areas for the 90 percent highest-density interval from the posterior distribution of parameters), model with high bank capital channel (dotted lines with circles), and model without imperfect interest rate pass-through (dashed lines with crosses).

The role of bank capital constraints and the macroeconomic propagation through the bank profit channel can also be illustrated in the case of a negative markup shock to mortgage loan spreads (figure 3), which creates an immediate boost to mortgage lending and consumption. It is observed the ensuing monetary tightening leads to a more pronounced reduction of output, and investment in particular, in the presence of a strong bank capital channel, as banks react to the increase in leverage by more forcefully raising their loan-deposit margin.

The cyclicality of bank profits and its impact on macroeconomic propagation is furthermore accentuated by the sluggishness of bank interest rate setting. A common finding in the empirical literature is that banks only gradually pass on the changes in monetary policy rates to the rates offered to their retail customers. This sluggishness may thus affect the speed and effectiveness of the monetary policy transmission via the interest rate channel. The frictions are furthermore often found to be asymmetric in the sense that bank lending rates tend to adjust quicker as a response to policy rate increases than to policy rate decreases.[21] In other words, the sluggishness of retail bank interest rates is another friction affecting the way shocks are propagated to the real economy. This is best illustrated when considering a monetary policy shock (figure 4). For example, a monetary tightening via its overall more muted impact on lending rates than on deposit rates, in turn, exerts an initial negative impact on bank capital that induces banks to adjust their loan-deposit margins to recoup their profits. This further amplifies the macroeconomic response to the initial monetary policy shock.

The importance of retail bank interest rate rigidities is furthermore considered for the case where banks have no market power when setting rates and where consequently the pass-through of policy rates to bank interest rates is immediate and complete are shown (dashed lines with crosses). Overall, this implies that monetary policy accommodation to the various shocks hitting the economy is transmitted fully and more quickly to the interest rates facing savers and borrowers. Hence, the counterbalancing impact of monetary policy is more powerful in this case. In other words, the common

[21] See, e.g., Mester and Saunders (1995), Mojon (2001), and Gropp, Kok Sørensen, and Lichtenberger (2007).

finding that the bank interest rate pass-through is sluggish implies a somewhat attenuated impact of the policy rate changes through the interest rate channel of monetary policy transmission.

4.2 Transitional Dynamics Towards Higher Capital Requirements

Our model is also well suited to investigate the macroeconomic implications of changes to the regulatory framework. The reform of the financial regulatory landscape enacted in end-2010 (so-called Basel III), following the proposal of the Basel Committee on Banking Supervision (BCBS), will lead to higher required capital for the banking sector.[22] The simulations presented remain illustrative of the transitional costs of introducing higher capital requirements but should not be interpreted as a quantitative economic assessment of the introduction of Basel III.[23] Indeed, the magnitude of the shock is not related to the exact calibration of the reform and to the balance sheet structure of the euro-area system. Moreover, the simulation is silent on the steady-state and cyclical benefits of higher capital requirements.

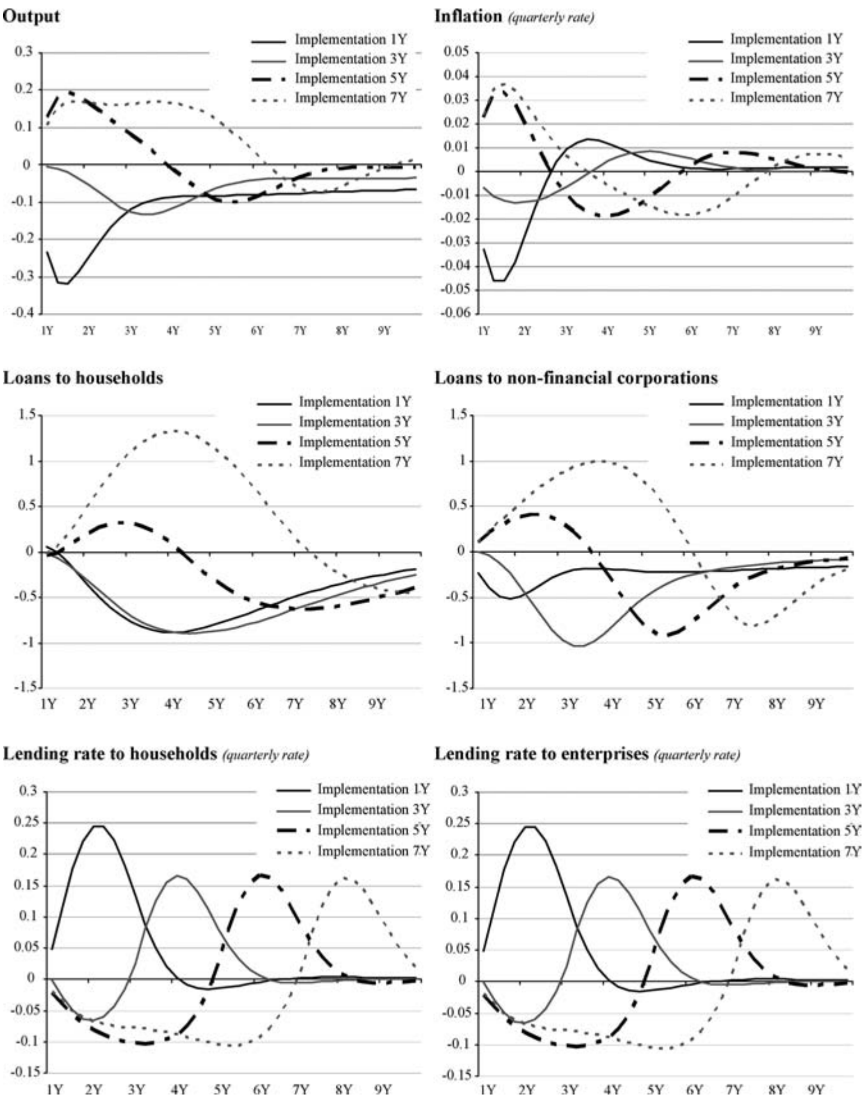
The simulations are presented in figure 5. As regards the timing of the introduction of the higher capital requirements, the experiments assume the implementation of higher capital requirements at different horizons (i.e., immediate implementation, after two, four, and six years, respectively). The model is run under perfect foresight and with endogenous monetary policy, following the estimated Taylor rule. Given the specification of the bank capital frictions and the calibration strategy for the steady state, capital requirements have no tangible impact on the real allocation over the long term. As described above, the required bank balance sheet adjustments take place through higher loan-deposit margins, which curb loan demand and support the internal capital accumulation through higher retained earnings.

The parameter driving the bank capital channel has been set at its highest value found across the various estimation exercises

[22]See Basel Committee on Banking Supervision (2011).

[23]For the macroeconomic impact assessment of introducing the new regulatory framework, see Basel Committee on Banking Supervision (2010).

Figure 5. Transitional Dynamics to Higher Capital Requirement for Different Implementation Dates: Benchmark Model



($\chi_{wb} = 50$). We also experimented with simulations through unexpected capital requirement shocks. This led to somewhat stronger effects which could even be more pronounced by assuming unchanged monetary policy. On balance, the perfect-foresight simulations presented below may be seen as the mid-range effects given possible assumptions on expectations and monetary policy reaction.

In the case of immediate implementation of higher capital requirements (solid black lines in figure 5), the maximum impact on real GDP is obtained after a few quarters. A 2-percentage-point increase in capital requirements leads to a peak decline in real GDP of 0.3 percentage point, the negative effects being rapidly reabsorbed over the medium term. The downward pressures on inflation are relatively short lived, reaching -0.05 percentage point of quarterly inflation after a few quarters and then reverting back to positive territory. As mentioned before, in the long term, the transition towards higher capital requirements leaves the real economy and the outstanding amount of loans unchanged since the adjustment will be fully reflected in higher bank capital. The required increase in bank profits depends on the magnitude of loan-deposit margins' increase compared with loan volume contraction. Figure 5 shows the hump-shaped responses of spreads and loans with opposite signs. Given the more gradual interest rate pass-through on mortgage lending rates, the price and volume adjustments of household credit are more sluggish than in the case of non-financial corporations.

Considering now the announcement of higher capital requirements at more distant horizons, it turns out that the output cost of bank balance sheet consolidation becomes smaller the later the implementation date. For implementation after three years, the peak negative impact on GDP is much more moderate and materializes later than in the previous case. The transition path of GDP even turns positive when higher capital requirements are expected to be implemented after five and seven years, respectively. In the latter cases, GDP only falls below baseline around the year of the implementation. The expansion of GDP in the first years is notably supported by lower bank lending rate spreads and is mainly caused by the forward-looking behavior of economic agents. In other words, borrowers decide to "front load" consumption and investment in view of expected tighter credit conditions in the future. This more benign impact on activity the further into the future the actual implementation of the new requirements is moved can be interpreted as a "smoothing out" of the negative implications of the capital shock. If banks have more time to adjust their activities and balance sheets to the new environment, they will tend to smooth the impact of the shock. The tighter the implementation schedule, the more important non-linearities in credit frictions will be.

4.3 Macroeconomic Propagation under Risk-Sensitive Capital Requirements

Not only has the level of required capital been increased under the Basel III agreement, but another key innovation already being implemented under the existing Basel II framework was the introduction of more risk-sensitive capital requirements. The regulatory interface of our model also allows for analyzing the macroeconomic propagation effects of such a change to the regulatory framework; i.e., moving from "fixed-rate" capital requirements (à la Basel I) to having risk-sensitive risk weights as stipulated under the Basel II and III agreements.

Under the risk-sensitive Basel II-like capital requirement regime, the static profit maximization problem of the bank is as follows:

$$\max_{B_t^w, Dep_t^w} R_{HH,t}^{wb} B_{HH,t}^{wb} + R_{E,t}^{wb} B_{E,t}^{wb} - R_t Dep_t^{wb} - \frac{\chi_{wb}}{2}(RWCap_t - 0.11)^2 Bankcap_t,$$

where

$$RWCap_t = \frac{Bankcap_t}{\left(a_0^E + a_1^E LEV_{E,t}^{wb} + b^E \varepsilon_{E,t}^{\sigma}\right) B_{E,t}^{wb} + \left(a_0^{HH} + a_1^{HH} LEV_{HH,t}^{wb} + b^{HH} \varepsilon_{HH,t}^{\sigma}\right) B_{HH,t}^{wb}}$$

and subject to the balance sheet identity

$$B_{HH,t}^{wb} + B_{E,t}^{wb} = Dep_t^{wb} + Bankcap_t.$$

$LEV_{E,t}^{wb}$ and $LEV_{HH,t}^{wb}$ are leverage ratios for the corporate and household sectors defined as debt over collateralized assets. a_0^E, a_1^E, b^E and $a_0^{HH}, a_1^{HH}, b^{HH}$ represent coefficients in the linearized version of the Basel II formula (see below for details). This formulation leads to the following lending spreads conditioned on the risk-sensitive capital requirements:

$$R^{wb}_{HH,t} - R_t = -\chi_{wb}(RWCap_t - 0.11) \times RWCap^2_t\big(a^{HH}_0 + 2a^{HH}_1 LEV^{wb}_{HH,t}\big)$$

$$R^{wb}_{E,t} - R_t = -\chi_{wb}(RWCap_t - 0.11)RWCap^2_t\big(a^E_0 + 2a^E_1 LEV^{wb}_{E,t}\big).$$

In contrast to the lending spreads derived under the Basel I regulatory regime, the target capital ratio is now dependent on the riskiness of the banks' borrowers, which is dependent on the state of the economy impinging on borrower net worth (via income and housing wealth on the side of households and via the value of the capital stock on the side of corporations).

For calculating the steady-state linear relationship between Basel II risk weights and leverage, we take as a starting point the Basel II risk-weight formulas and subsequently linearize the resulting risk curves for entrepreneurs and households around their respective steady-state leverage ratios.

As a first step, under the Basel II capital adequacy framework, the risk-weighted assets are derived using the following formulas.[24] The capital requirement formula for the corporate exposures is given by

$$CR^E = LGD^E\Phi\left[(1-\tau^E)^{-0.5}\Phi^{-1}PD^E + \left(\frac{\tau^E}{1-\tau^E}\right)^{0.5}\Phi^{-1}(0.999)\right] - PD^E LGD^E,$$

where PD^E and LGD^E refer to probability of default and loss given default on corporate exposures, respectively. Φ denotes the cumulative distribution function for a standard normal random variable.

[24] We focus here on the foundation internal ratings-based approach and assume fixed LGD values provided by the supervisory authority. For corporate exposures (i.e., entrepreneurs) we assume an LGD value of 0.45 and for household exposures we assume an LGD value of 0.35 (retail mortgage exposures are presumably better collateralized, hence the lower LGD). We furthermore, for simplicity, assume a one-year maturity. For more details on the Basel II formulas, see Basel Committee on Banking Supervision (2004).

τ^E denotes the asset-value correlation which parameterizes cross-borrower dependencies and, being a decreasing function of PD, is equal to

$$\tau^E = 0.12\left[\frac{(1-\exp(-50PD^E))}{(1-\exp(-50))}\right] + 0.24\left[1 - \frac{(1-\exp(-50PD^E))}{(1-\exp(-50))}\right].$$

As we assume a fixed LGD (equal to 0.45), the only time-varying component in the risk weighting is the PD, and the resulting risk curve has a concave nature.

For household exposures, we apply the following derivation of the capital requirement:

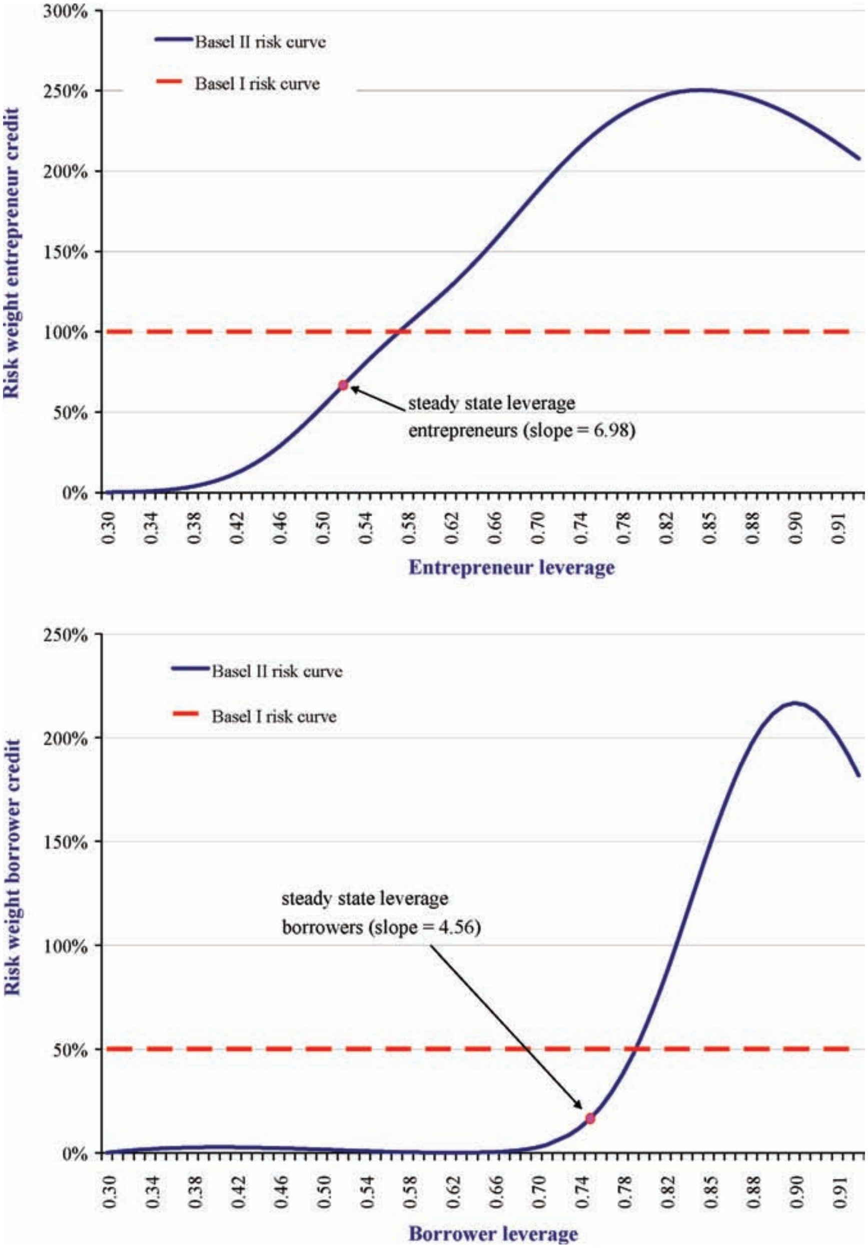
$$CR^{HH} = LGD^{HH}\Phi\left[(1-\tau^{HH})^{-0.5}\Phi^{-1}PD^{HH} + \left(\frac{\tau^{HH}}{1-\tau^{HH}}\right)^{0.5}\Phi^{-1}(0.999)\right] - PD^{HH}LGD^{HH},$$

where τ^{HH} equals 0.15. Also in the case of household exposures, the time variation of the risk curve is a function of PDs only (as LGD^{HH} is fixed at 0.35). The risk-weighted assets are subsequently derived as $RWA^E = CR^E * 12.5 * 1.06 * EAD^E$ and $RWA^{HH} = CR^{HH} * 12.5 * EAD^{HH}$, where EAD denotes exposure at default (i.e., $B^{wb}_{E,t}$ and $B^{wb}_{HH,t}$ for corporate exposures and household exposures, respectively).[25] The time-varying correlation adjustment parameter and the assumed higher LGD for corporate exposures results in higher risk weights and an initially steeper risk curve relative to the risk function with respect to household exposures.

In the next step, the Basel II-based risk-weight functions can be expressed in terms of borrower leverage, $(G(\varpi))$ for households and $(G_E(\varpi_E))$ for entrepreneurs. As can be seen from figure 6, there is a positive and concave relationship between required capital and the leverage of borrowers, which in turn is a positive function of the probability of default, $(\overline{\varpi}_{HH,t})$ and $(\overline{\varpi}_{E,t})$ for households and entrepreneurs, respectively.

[25]The scaling factor of 1.06 in the calculation of the risk-weight function for corporate exposures aims at compensating for the expected overall decline in capital requirements caused by the transition from Basel I to Basel II.

Figure 6. Risk Weights under Basel I and Basel II



Mechanically, owing to the risk-weight functions, it can be conjectured that shocks to borrower credit risk would give rise to higher capital requirements. As credit risk often deteriorates in economic downturns and improves in upturns, it has been argued that the regulatory risk curves as formulated in Basel II could have amplifying procyclical effects on the business cycle (to the extent that bank capital constrains bank lending, which in turn may be an imperfect substitute to other financing sources).[26] At the same time, if banks engage in active management of their loan portfolio, either as a response to or in anticipation of cyclical requirements to their minimum capital levels, the overall effect on the business cycle may not be as mechanical as what the simple transposition of the risk weighting to capital requirements and lending would prescribe.[27]

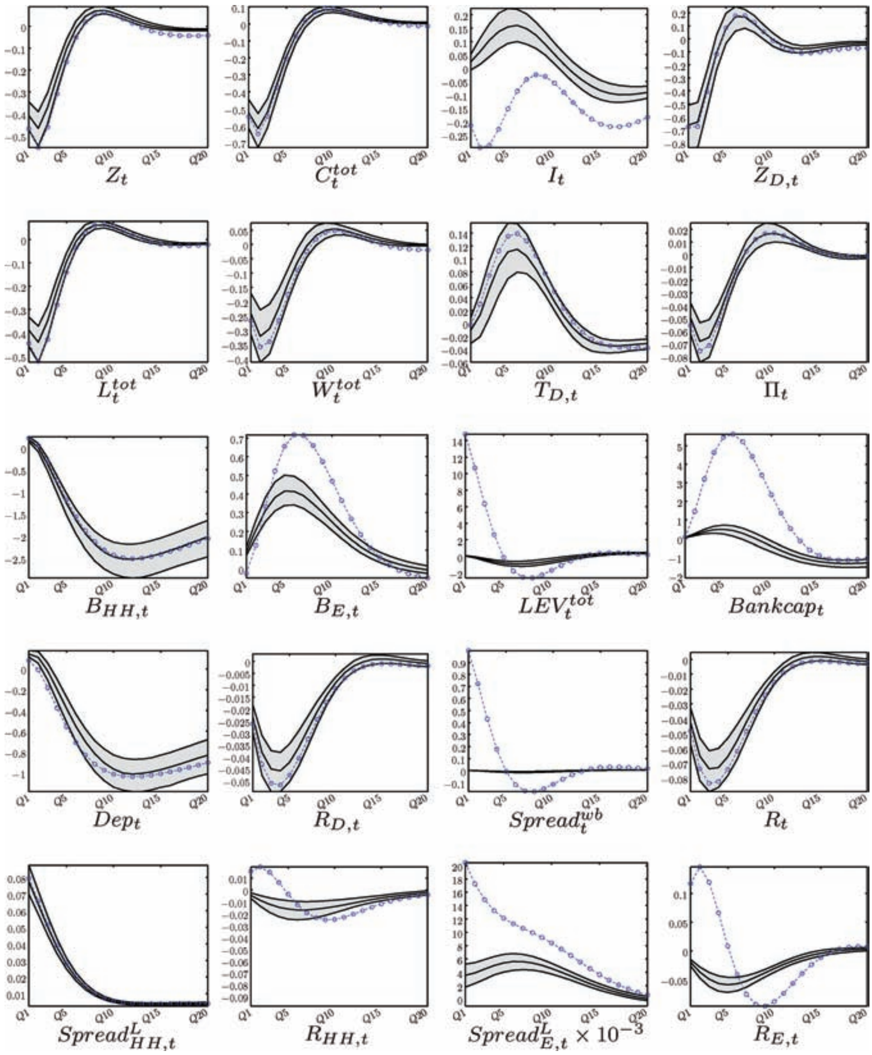
The DSGE model has been estimated on euro-area data, assuming constant capital requirements over the cycle, which is interpreted as consistent with Basel I regulatory framework. Given the estimated sources of business-cycle fluctuations, we simulate a counterfactual economy where capital requirements are risk sensitive according to the Basel II risk-weights formula. The model considers two types of risky assets: loans to households for house purchase and loans to non-financial corporations. The counterfactual economy under Basel II turns out to be marginally more volatile overall, with unchanged monetary policy rule. Compared with economic fluctuations under Basel I, risk-sensitive capital requirements imply 5 percent higher volatility in real GDP growth and 4 percent higher volatility in inflation.

The relatively limited impact on macroeconomic volatility masks more pronounced amplification mechanisms for specific sources of economic disturbances, and notably financial shocks. Figures 7 and 8 illustrate the impact of more risk-sensitive capital requirements on real and financial variables. Focusing on the different shock amplifications in the benchmark model (i.e., Basel I; solid black lines) and the Basel II-based benchmark model (dotted lines with circles), we observe that, for example, a shock to borrower riskiness has a

[26] See, e.g., Danielsson et al. (2001), Kashyap and Stein (2004), and Catarineau-Rabell, Jackson, and Tsomocos (2005).

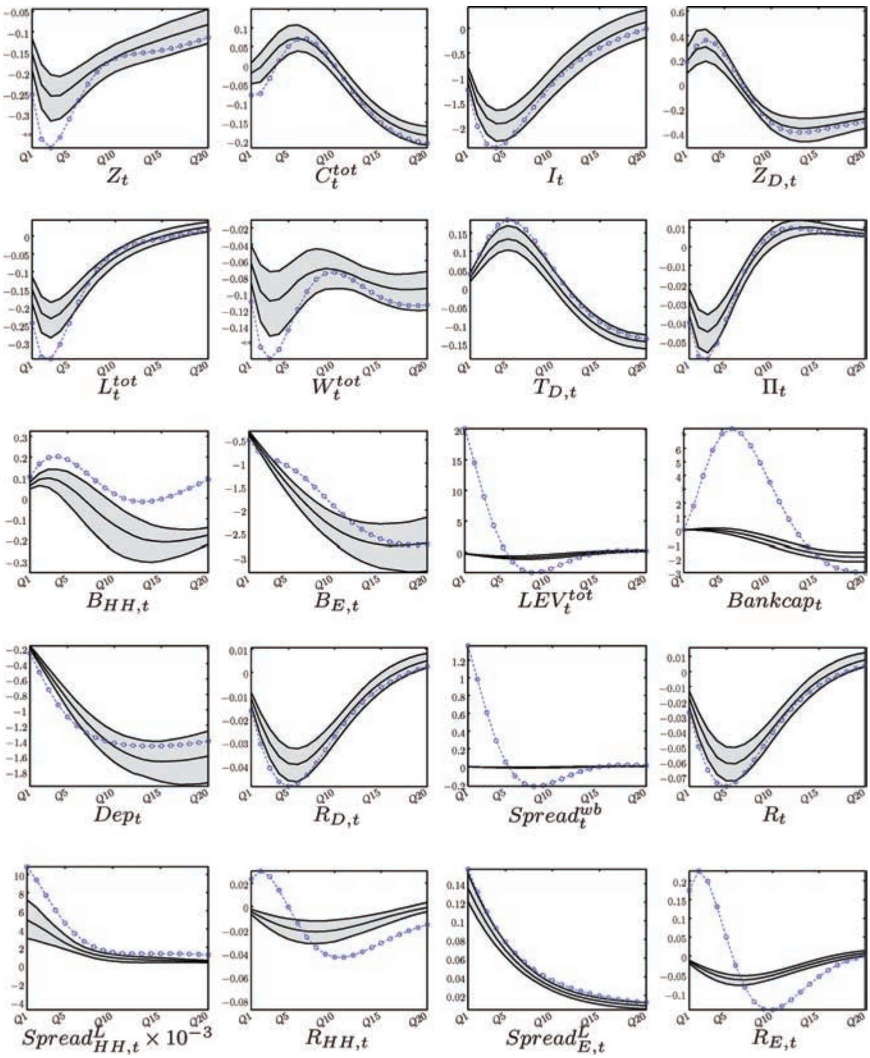
[27] See, e.g., Gordy and Howells (2006), Zhu (2008), Boissay and Kok Sørensen (2009), and Jokivuolle, Kiema, and Vesala (2009).

Figure 7. Impulse Response Functions Associated with a Shock on $\varepsilon^{\sigma}_{HH,t}$



Note: Benchmark model (plain lines for the mode and shaded areas for the 90 percent highest-density interval from the posterior distribution of parameters), and benchmark model under Basel II (dotted lines with circles).

Figure 8. Impulse Response Functions Associated with a Shock on $\varepsilon^{\sigma}_{E,t}$



Note: Benchmark model (plain lines for the mode and shaded areas for the 90 percent highest-density interval from the posterior distribution of parameters), and benchmark model under Basel II (dotted lines with circles).

more pronounced impact on lending spreads when banks are subject to risk-sensitive capital requirements. In contrast to the benchmark case, bank lending rates increase, allowing banks to rebuild their capital in response to the higher (risk-weighted) leverage. In the case of a negative shock to corporate riskiness, investment is more adversely affected under the Basel II framework, and the positive spillover impact on consumption is more muted relative to the baseline (Basel I). Likewise, in the case of an adverse shock to household default risk, the need for banks to accumulate more capital results in a negative spillover effect on the corporate sector (via higher corporate lending spreads). Overall, we observe that changes in credit risk across time, especially in the case of a shock to corporate creditworthiness, amplifies the impact on output compared to the situation with flat-rate capital requirements. This underlines the importance of banks' risk perception in guiding their lending behavior and stresses its potential amplifying effect on economic fluctuations. Sharp deteriorations in the creditworthiness of households and firms—as, for example, observed during the 2007–09 financial crisis[28]—are therefore likely to produce reverberating feedback effects on real economic activity.

This notwithstanding, it is notable that under risk-sensitive capital requirements, banks are found to more actively reshuffle their loan portfolio in response to credit-risk shocks—as, for example, illustrated by the stronger reaction of the volumes of corporate loans and mortgage loans to a shock to household and corporate creditworthiness, respectively. This might hence exert a mitigating impact on the procyclical nature of the risk-sensitive capital requirements, although in our specification it is not enough to completely eliminate the cyclical propagation mechanism of the Basel II framework.

[28]For example, expected default frequencies of euro-area non-financial corporations (which is a measure of corporate default risk produced by Moody's KMV) increased sixfold between June 2007 and December 2009. Likewise, according to the ECB Bank Lending Survey, the net percentage of banks reporting that risk perceptions contributed to a tightening of credit standards increased from 9 percent in 2007:Q2 to 46 percent in 2008:Q4 with respect to mortgage loans and from −4 percent in 2007:Q2 to 64 percent in 2008:Q4 with respect to loans to enterprises.

4.4 Accounting for Countercyclical Macroprudential Policies

A final application of the model is devoted to the interactions between monetary policy and macroprudential policy. In particular, we want to assess whether a countercyclical regulatory regime would support macroeconomic stabilization. Recent papers like Angeloni and Faia (2009) or Kannan, Rabanal, and Scott (2009) have investigated this issue with different formulations of the strategic interactions between monetary policy and macroprudential policy. Here we focus on the joint determination of the two policy rules so as to maximize an ad hoc loss function under credible commitment.

The intertemporal quadratic loss function penalizes deviations from steady state for consumer price inflation, output growth, and policy rate. Monetary policy conduct is described as an interest rate rule, while macroprudential policy is assumed to follow a capital requirement rule. Both rules feature policy inertia and respond to level and first difference of consumer inflation, detrended output, and first difference of loans to households, loans to entrepreneurs, real housing prices, and real equity prices.[29] We chose to limit the analysis to a stylized loss function instead of a welfare-based objective, as the "reduced-form" nature of the bank capital friction considered in this paper would weakly portray the welfare trade-offs faced by macroprudential policy in particular. Consequently, we preferred to abstract from welfare calculations and gear the policy discussion towards general macroeconomic stabilization without investigating how the microfoundations of the model influence the policy objectives.

The loss function considered can be written as follows:

$$\mathcal{L}_t = \lambda_\pi \pi_t^2 + \lambda_z [\Delta z_t]^2 + \lambda_r r_t^2 + \lambda_{lev} [Leverage_t]^2 + \beta \mathbb{E}_t \mathcal{L}_{t+1},$$

where λ_π, λ_z, and λ_r are the coefficients weighting the respective costs of volatility in CPI inflation, changes in output, and nominal interest rate. Later on, we would consider introducing a penalty for bank leverage volatility.

The weights in the loss function are selected in the following way. The monetary policy rule has the same form as the estimated one.

[29]Real equity prices are defined as the average real price of fixed capital in the economy.

The exogenous processes for the structural shocks are taken from the benchmark estimation. Then we search for the weighting scheme which delivers, at the optimal rule, the same volatility for inflation and policy rate as under the estimated rule. The optimal weights we obtain are $\lambda_\pi = 1$, $\lambda_z = 4$, and $\lambda_r = 0.75$. Such a loss function constitutes an intuitive benchmark. Another possibility would have been to consider the full efficiency curve in the inflation, output growth space. But, for the sake of clarity, we kept only one specific loss function. The essence of the results presented thereafter holds for any point of this efficiency curve.

A first exercise consists of optimizing the parameters of the monetary policy rule augmented with asset prices and credit variables, keeping capital requirements constant. We concentrate on the following formulation of the monetary policy rule:

$$r_t = \rho r_{t-1} + (1-\rho)(r_\pi \pi_{t-1} + r_y y_{t-1}) + r_{\Delta\pi}\Delta\pi_t + r_{\Delta y}\Delta y_t \\ + r_{T_D}\Delta t_{D,t} + r_Q \Delta q_t + r_h \Delta b_{HH,t} + r_e \Delta b_{E,t}.$$

We only consider financial shocks, as provided by the benchmark estimation: those disturbances relate to interest rate markups, borrowers' risk, bank capital, and housing preference, which also introduced its contribution to housing prices. Focusing on economic disturbances at the core of credit intermediation enables us to present more striking results on the role of credit and asset prices for monetary policy conduct in interaction with a countercyclical regulatory framework. As sensitivity analysis (not presented here), we verified that the findings exposed thereafter were still holding when all shocks were introduced.

Table 4 presents the macroeconomic volatilities associated with various optimized rules in the presence of financial shocks (except for the first column). In the first two columns, the monetary policy rule is specified as in the estimation and optimized under constant capital requirements. For the sake of completeness, the exercise is conducted either with financial shocks or with the overall set of economic disturbances. In both cases, the optimized monetary policy rule features a high level of interest rate inertia, a strong long-term response to inflation, stronger reaction to changes in output than in its level, and a specific role for housing prices. The restriction to financial shocks seems to increase the coefficient on housing prices

Table 4. Optimized Monetary and Macroprudential Policy Rules

Loss Function						
λ_π				1		1
λ_z				4		4
λ_r				0.75		0.75
λ_{lev}				0		0.0001
Regulatory Regime	*Basel I All Shocks*	Basel I *Bench.*	Basel I	Basel II	Countercyclical	Countercyclical
Optimized Policy Parameters						
ρ	*0.98*	0.95	0.96	0.93	0.997	0.96
r_π	*43.90*	43.90	52.06	52.07	43.91	43.91
$r_{\Delta\pi}$	*0.53*	0.75	1.12	1.13	−0.43	0.43
r_y	*0.57*	0.75	0.04	0.07	0.92	0.93
$r_{\Delta y}$	*0.56*	1.74	2.30	2.24	1.61	1.99
r_{T_D}	*0.20*	0.68	0.41	0.63	0.00	0.26
$r_{\Delta h}$	—	—	0.45	0.63	0.00	0.36
r_Q	—	—	0.01	0.07	0.00	0.00
$r_{\Delta e}$	—	—	−0.08	−0.12	0.00	0.02
ρ^{bc}	—	—	—	—	0.78	0.77
r_y^{bc}	—	—	—	—	113.00	0.00
$r_{\Delta y}^{bc}$	—	—	—	—	0.40	0.00
$r_{T_D}^{bc}$	—	—	—	—	0.03	0.13
$r_{\Delta h}^{bc}$	—	—	—	—	−0.05	0.01
r_Q^{bc}	—	—	—	—	−1.91	−0.38
$r_{\Delta e}^{bc}$	—	—	—	—	−0.43	−0.11

(continued)

Table 4. (Continued)

Loss Function						
λ_π				1		1
λ_z				4		4
λ_r				0.75		0.75
λ_{lev}				0		0.0001
Regulatory Regime	*Basel I All Shocks*	Basel I *Bench.*	Basel I	Basel II	Countercyclical	Countercyclical
Relative STD to Bench. (in %)						
ΔZ_t	—	100.0	80.3	102.3	16.5	78.6
Π_t	—	100.0	139.8	116.7	71.6	138.2
R_t	—	100.0	72.0	91.9	29.7	65.1
$T_{D,t}$	—	100.0	100.0	96.1	104.6	100.6
$B_{HH,t}$	—	100.0	97.0	84.7	227.8	103.2
$B_{E,t}$	—	100.0	99.9	80.4	136.8	94.4
$Leverage_t$	—	100.0	99.0	230.1	482.4	94.6
$\mathcal{L}$	—	0.34	0.23	0.40	0.03	0.32

and output growth but does not change qualitatively the main properties on the monetary policy rule. The macroeconomic variances generated by this monetary policy rule are taken as benchmark to normalize the moments obtained with the other policy regimes in table 4.

In the third column, we allow for monetary policy reaction to credit and equity prices. The augmented optimal rule improves upon the previous one, reducing the loss function from 0.34 to 0.23. However, the lower volatility obtained for output growth and the interest rate is counterbalanced by a higher standard deviation for inflation. This optimal rule still displays a high degree of interest rate inertia, a strong reaction to inflation, and some specific role for housing prices. But in addition, the rule includes some positive response to household loans, whereas the coefficients on loans to entrepreneurs and real equity prices are close to zero. Even without introducing asset prices or credit in the objective function, it turns out that the financial frictions on the household side vindicate some specific monetary policy focus on credit and asset prices.

With the augmented monetary policy rule specification, we also investigated the implications of risk-sensitive capital requirements. In this case, the optimized coefficients remain very close to the ones obtained with constant capital requirements (see column 4 in table 4). At the margin, the monetary policy response to housing prices and household loans turns out to be stronger.

In the last two columns of table 4, we allow for time-varying capital requirements. We assume that the target bank capital ratio follows a log-linear rule of the form

$$cap_t = \rho^{bc} cap_{t-1} + r_y^{bc} y_t + r_{\Delta y}^{bc} \Delta y_t + r_{\Delta h}^{bc} \Delta b_{HH,t} + r_{\Delta e}^{bc} \Delta b_{E,t} + r_{T_D}^{bc} \Delta t_{D,t} + r_Q^{bc} \Delta q_t.$$

Keeping the same loss function as in the previous experiments, the joint optimal determination of policy rules suggests that countercyclical regulation could provide a strong support to macroeconomic stabilization. The optimized capital requirement rule features some inertia and a very high positive response to output, while the role for credit variables and asset prices seems negligible. The optimized monetary policy rule is very much affected by the introduction of countercyclical regulation: in particular, all coefficients on credit and asset prices become insignificant. Acting at the core of the financial system, regulatory policy seems to be relatively more effective than monetary policy in addressing destabilizing fluctuations in credit markets and intratemporal wedges between financial costs, therefore alleviating somehow the need for monetary policy to "lean against the wind." The jointly determined policy rules deliver a superior macroeconomic outcome. The loss function gets close to zero, with output growth volatility at 16.5 percent of the benchmark, inflation volatility at 70 percent, and interest rate at 30 percent. However, in the model, the main transmission channel of regulatory policy on the economy works through the adjustment of bank balance sheets and its impact on bank lending rates. Consequently, the macroeconomic stabilization support from the optimized capital requirement rule implies an almost fivefold increase in bank leverage volatility. Such a degree of countercyclical capital requirements would therefore be difficult to implement and lead to excessive volatility in bank balance sheets. As shown in the last column of table 4, if we constrain the regulatory framework by introducing a relatively small penalty for leverage volatility in the loss function, then the optimized capital requirement rule becomes only moderately time varying and the monetary policy rule is very similar to the one obtained under constant capital requirements.

Overall, while some countercyclical regulation seems suitable as far as macroeconomic stabilization is concerned, its design and magnitude should be carefully considered. The analysis presented here remains illustrative and subject to clear limitations. Notably, a structural interpretation of systemic risk (and in particular its cross-sectional dimension) is absent from the model. Such a concept is essential to define a meaningful objective for macroprudential policy.

5. Conclusions

The recent years' dramatic events which brought financial markets into turmoil highlighted the crucial role of credit market frictions in the propagation of economic and financial shocks. However, the nature of banking and the role of banks in amplifying macroeconomic fluctuations are elements that hitherto have been largely neglected in the macroeconomic literature and, in particular, in the design of general equilibrium models. To reflect this, a number of recent papers try to correct this void by incorporating banking sectors and other financial frictions into DSGE modeling frameworks. The model presented in this paper contributes to this research by incorporating a number of demand and supply credit frictions into an estimated DSGE model of the euro area.

Apart from documenting the potential amplifying effects of credit frictions, this setup allows us to analyze changes in the regulatory regimes facing the financial sector, such as the introduction of risk-sensitive capital requirements or the transition towards more stringent regulatory regimes. Moreover, reflecting the renewed focus on the nexus between monetary policy and macroprudential (or financial-stability-oriented) policies, our results point to important complementarities.

Finally, a few caveats and directions for further research should be mentioned. First of all, the banking sector in our setup is of a reduced-form nature and can be further improved. For example, a more complete description of the balance sheet composition of the banks taking into account issues such as liquidity, wholesale funding, and trading book valuations would enhance the specification and also allow for analyzing the macroeconomic impact of money market disruptions, bank liquidity positions, and unconventional monetary policies. Likewise, a more microfounded optimization of the policy rule to study the interactions between macroprudential and monetary policies could be pursued.

Appendix 1. Data

Data for GDP, consumption, investment, employment, wages, and consumption deflator are taken from Fagan, Henry, and Mestre (2001) and Eurostat. Employment numbers replace hours. Consequently, as in Smets and Wouters (2005), hours are linked to the number of people employed, e_t^*, with the following dynamics:

$$e_t^* = \beta \mathbb{E}_t e_{t+1}^* + \frac{(1-\beta\lambda_e)(1-\lambda_e)}{\lambda_e}\left(l_t^* - e_t^*\right).$$

House prices for the euro area are based on national sources and taken from the ECB web site.[30] Residential investment is taken from Eurostat national accounts and is backdated using national sources. Households' debt for the euro area also comes from the ECB and Eurostat.[31] The three-month money market rate is the three-month Euribor taken from the ECB web site, and we use backdated series for the period prior to 1999 based on national data sources. Household deposits are proxied using a backdated series of M2 which is available from the ECB web site and which represents the main part of deposits held with monetary financial institutions (MFIs) by euro-area non-financial private-sector residents (households primarily). Data on MFI loans to households and non-financial corporations are likewise taken from the ECB web site. Data prior to September 1997 have been backdated based on national sources. Meanwhile, data on retail bank loan and deposit rates are based on official ECB statistics from January 2003 onwards and on ECB internal estimates based on national sources in the period before. The lending rates refer to new business rates on loans to households for house purchase and new business rates on loans to non-financial corporations, excluding bank overdrafts. For the period prior to January 2003, the euro-area aggregate series have been weighted using corresponding loan volumes (outstanding amounts) by country. Deposit rates refer to MFI interest rates on time deposits with agreed maturity taken from households. Similar to the derivation of the loan rates, from January 2003 deposit rates are based on official ECB statistics; prior to that period, they are based on a volume-weighted average of country-based rates.

Appendix 2. Calibrated Parameters and Steady State

Some parameters are excluded from the estimation and have to be calibrated. These are typically parameters driving the steady-state values of the state variables, for which the econometric model based on detrended data is almost non-informative.

[30] We applied some statistical interpolation methods to generate quarterly series.

[31] See the ECB's *Monthly Bulletin*, October 2007, for the description of the data used.

The discount factors are calibrated to 0.995 for the patient agents and 0.96 for the impatient agents and entrepreneurs.[32] The implied equilibrium real deposit interest rate is 2 percent in annual terms.[33] The depreciation rate for housing, δ, is equal to 0.01, corresponding to an annual rate of 4 percent, whereas the depreciation rate of capital, δ_X, is set to 0.1. Markups are equal to 1.3 in the goods markets (for both non-residential and residential goods) and 1.5 in the labor market (in each sector). The relative share of residential goods in the utility function, ω_D, is set to 0.1 for both household types. The value is chosen to pin down the steady-state ratio of residential investment to GDP. The intratemporal elasticity of substitution, η_D, is equal to 1. The intertemporal elasticity of substitution of entrepreneurs is set to 1 (σ_{CE}). The relative shares of inputs in production are 0.3 for capital (α) and 0.7 for labor in the non-residential goods sector, while in the residential sector we assign a weight equal to 0.1 to land ($\alpha_{\mathcal{L}}$) and reduce the share of capital to 0.2 (α_D), in order to maintain the level of labor intensity unchanged.

The markups on loan and deposit rates are calibrated so that the margin between the loan rate and the deposit rate is 100 basis points in annual terms, while the annual spreads on lending rates to households and entrepreneurs are 200 basis points and 120 basis points, respectively. Those numbers are very close to the historical averages from 1999:Q1 to 2008:Q2.[34] Given the discount factors and the markups on retail interest rates, the steady-state values for the default cut-off points $\overline{\varpi}_E, \overline{\varpi}_{HH}$ are numerically determined by the modified Euler equations of borrowers and entrepreneurs. Once those cut-off points are computed—and assuming monitoring costs of 0.2 for non-financial corporations, μ_E, and 0.15 for households, μ_{HH}—the standard deviations of the idiosyncratic shocks are adjusted to

[32] See, e.g., Iacoviello (2005), Monacelli (2009), and Iacoviello and Neri (2010) for a thorough discussion of the calibration of the discount factors in a similar setup.

[33] The steady-state level of the interest rate is pinned down by the savers' intertemporal discount factor.

[34] We confine the calibration of the loan-deposit margin and the lending spreads to the period starting in 1999:Q1, as due to the convergence of interest rates prior to the introduction of the euro there was a gradual downward level shift in loan and deposit rates in the years preceding 1999. Because of this structural shift in the level of rates, for the steady-state calibration we apply the pattern of loan and deposit rates for the euro period only.

reproduce default frequencies for impatient households and firms of 0.3 percent and 0.7 percent, respectively.[35]

Finally, we set in the benchmark estimation the share of borrowers ω at 0.25. The loan-to-value ratios (determined by the terms $(1-\chi_E)$ for non-financial corporations and $(1-\chi_{HH})$ for impatient households) are then determined to ensure plausible debt-to-GDP ratio in the steady state. With $(1-\chi_E)$ at 0.6 and $(1-\chi_{HH})$ at 0.2, the share of corporate loans to annual GDP is around 33 percent, while the share of household housing loans to annual GDP is around 25 percent. This calibration is close to the levels recorded in the euro area around the year 2000 as well as to their historical average levels since 1980. Besides, the loan-to-value (LTV) ratios are consistent with the available range of estimates.[36]

Appendix 3. Prior Distributions

The standard errors of the structural shocks are assumed to follow a uniform distribution, while the persistence parameters follow a beta distribution with mean 0.5 and standard deviation 0.2.

Regarding the parameters of the monetary policy reaction function, we follow Smets and Wouters (2005) quite closely. The interest rate smoothing parameter follows a beta distribution with parameters 0.75 and 0.1. The parameters capturing the response to changes in inflation and output gap follow a gamma distribution with parameters 0.3 and 0.1, and 0.12 and 0.05, respectively. Concerning the response to inflation and output gap, the prior distributions

[35]This is consistent with corporate default statistics from Moody's, the rating agency, which show an average default rate on (non-U.S.) non-financial corporate bonds of 0.75 percent for the period 1989–2009. Household default rates can be approximately derived using the loan write-off data in the ECB's MFI balance sheet statistics. Computing the ratio of average write-offs on mortgage loans to corporate loans for the period of available data (2001–09), it is found that the share of defaulting mortgage loans to corporate loans is c. 45 percent. Hence, using the non-financial corporate default rate derived from Moody's implies an approximate mortgage default rate of 0.34 percent, i.e., close to our steady-state calibrated value.

[36]LTV ratios for euro-area housing loans differ across countries but tend on average to lie in the range of 0.7–0.8 percent; see European Central Bank (2009). LTV ratios can be approximated by the debt-to-financial-asset ratio of the non-financial corporate sector, which on average between 1999 and 2009 was around 0.45 (sources: ECB and Eurostat and ECB calculations).

are a normal with mean 2.5 and standard deviation 0.25, and a gamma with parameters 0.12 and 0.05, respectively. The prior on the level inflation terms has been increased compared with the empirical DSGE literature, as the determinacy region in the two-sector economy considered in this paper requires stronger reaction to price pressures.

Regarding preference parameters, the intertemporal elasticity of substitution, which is common to both household types, follows a gamma distribution with mean 1.2 and standard deviation 0.2. The habit formation parameter is also the same for savers, borrowers, and entrepreneurs, following a beta distribution with parameters 0.75 and 0.1. The elasticity of labor supply is the same for both household types and sectors, and has a gamma(1.5, 0.1) prior distribution. On the production side, the adjustment cost parameters for fixed investment and the capacity utilization elasticity, which are common to both sectors, follow respectively a normal(4, 1.5) and a beta(0.5, 0.15) prior distribution. The prior distribution regarding the adjustment cost parameter for residential investments of savers and borrowers is a gamma(1,0.5). About nominal rigidities, the Calvo parameters for price setting in the non-residential sector and wage settings in each sector are distributed according to a beta distribution with mean 0.75 and standard deviation 0.05.[37] The indexation parameters are instead centered around 0.5, with a standard deviation of 0.15. In the residential sector, we set lower priors for the nominal price rigidities, with a beta(0.2,0.1) given assumptions made in the literature on the flexibility of housing prices (see Iacoviello and Neri 2010, for example). We do not introduce indexation on past inflation in the residential sector price setting.

Turning to the Calvo parameters driving the imperfect pass-through of policy rate on lending rates, we choose fairly uninformative priors with beta(0.5,0.2). The sensitivity of bank spreads on bank capital ratio inadequacy has relative tight priors, with a gamma(20,2.5), as in Gerali et al. (2010). Finally, in the benchmark model, the share of borrowers is not estimated, but in alternative

[37]In the estimation exercise we impose that the same level of nominal rigidity applies to the saver's and borrower's wages in a given sector. Such restriction is motivated by the availability of sector-specific, as opposed to individual-specific, data on wages.

specifications we introduce priors following beta distribution, with mean 0.35 and standard deviation 0.05. This choice is similar to the one of Iacoviello and Neri (2010). The model is still well defined when the share of borrowers goes to zero so that the estimation of the parameters is not affected by a singular point in zero.

References

Adrian, T., and H. S. Shin. 2009. "Liquidity and Leverage." *Journal of Financial Intermediation* 19 (3): 418–37.

Agenor, P.-R., and K. Alper. 2009. "Monetary Shocks and Central Bank Liquidity with Credit Market Imperfections." Discussion Paper No. 120, Centre for Growth and Business Cycles Research.

Agenor, P.-R., and L. A. Pereira da Silva. 2009. "Cyclical Effects of Bank Capital Requirements with Imperfect Credit Markets." Policy Research Working Paper No. 5067, World Bank.

Aguiar, A., and I. Drumond. 2009. "Business Cycle and Bank Capital Requirements: Monetary Policy Transmission under the Basel Accords." Manuscript, CEMPRE (March).

Angeloni, I., and E. Faia. 2009. "A Tale of Two Policies: Prudential Regulation and Monetary Policy with Fragile Banks." Working Paper No. 1569, Kiel Institute for the World Economy.

Ayuso, J., D. Pérez, and J. Saurina. 2004. "Are Capital Buffers Pro-cyclical? Evidence from Spanish Panel Data." *Journal of Financial Intermediation* 13 (2): 249–64.

Basel Committee on Banking Supervision. 2004. *International Convergence of Capital Measurement and Capital Standards — A Revised Framework*. Basel, Switzerland: Bank for International Settlements.

———. 2006. "Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework — Comprehensive Version." (June).

Basel Committee on Banking Supervision. 2010. "Assessing the Macroeconomic Impact of the Transition to Stronger Capital and Liquidity Requirements — Final Report." Bank for International Settlements (December).

Basel Committee on Banking Supervision. 2011. *Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems*. Basel, Switzerland: Bank for International Settlements.

Berger, A. N., R. DeYoung, M. J. Flannery, D. K. Lee, and O. Oztekin. 2008. “How Do Large Banking Organizations Manage Their Capital Ratio?” Research Working Paper No. 08-01, Federal Reserve Bank of Kansas.

Bernanke, B. S., M. Gertler, and M. Gilchrist. 1999. “The Financial Accelerator in a Quantitative Business Cycle Framework.” In *Handbook of Macroeconomics*, ed. J. B. Taylor and M. Woodford. North-Holland.

Bernanke, B. S., and C. Lown. 1991. “The Credit Crunch.” *Brookings Papers on Economic Activity* 2: 205–48.

Bikker, J., and P. Metzemakers. 2004. “Is Bank Capital Procyclical? A Cross-Country Analysis.” Working Paper No. 009, De Nederlandsche Bank.

Boissay, F., and C. Kok Sørensen. 2009. “The Stabilizing Effects of Risk-Sensitive Bank Capital.” Working Paper, European Central Bank.

Calvo, G. 1983. “Staggered Prices in a Utility-Maximizing Framework.” *Journal of Monetary Economics* 12 (3): 383–98.

Cappiello, L., A. Kadareja, C. Kok Sørensen, and M. Protopapa. 2010. “Do Bank Loans and Credit Standards Have an Effect on Output? A Panel Approach for the Euro Area.” ECB Working Paper No. 1150.

Carlstrom, C. T., and T. S. Fuerst. 1997. “Agency Costs, Net Worth, and Business Fluctuations: A Comparable General Equilibrium Analysis.” *American Economic Review* 87 (5): 893–910.

Catarineau-Rabell, E., P. Jackson, and D. Tsomocos. 2005. “Procyclicality and the New Basel Accord — Banks’ Choice of Loan Rating System.” *Economic Theory* 26 (3): 537–57.

Christensen, I., and A. Dib. 2008. “The Financial Accelerator in an Estimated New Keynesian Model.” *Review of Economic Dynamics* 11 (1): 155–78.

Christiano, L., R. Motto, and M. Rostagno. 2007. “Shocks, Structures, or Monetary Policies? The Euro Area and U.S. after 2001.” NBER Working Paper No. 13521.

———. 2010. “Financial Factors in Economic Fluctuations.” ECB Working Paper No. 1192.

Ciccarelli, M., A. Maddaloni, and J. L. Peydró. 2009. “Trusting the Bankers: A New Look at the Credit Channel of Monetary

Transmission." Paper presented at the C.R.E.D.I.T conference, Venice, September 24–25.

Covas, F., and S. Fujita. 2009. "Procyclicality of Capital Requirements in a General Equilibrium Model of Liquidity Dependence." Working Paper 09-23, Federal Reserve Bank of Philadelphia (September).

Danielsson, J., P. Embrechts, C. Goodhart, C. Keating, F. Muennich, O. Renault, and H. S. Shin. 2001. "An Academic Response to Basel II." Special Paper No. 130, LSE Financial Markets Group.

Darracq Pariès, M., and A. Notarpietro. 2008. "Monetary Policy and Housing Prices in an Estimated DSGE Model for the U.S. and the Euro Area." ECB Working Paper No. 972.

Darracq Pariès, M., C. Kok Sørensen, and D. Rodriguez Palenzuela. 2010. "Macroeconomic Propagation under Different Regulatory Regimes: Evidence from an Estimated DSGE Model for the Euro Area." ECB Working Paper No. 1251.

de Bondt, G. J. 2005. "Interest Rate Pass-Through: Empirical Results for the Euro Area." *German Economic Review* 6 (1): 37–78.

de Walque, G., O. Pierrard, and A. Rouabah. 2009. "Financial (In)Stability, Supervision and Liquidity Injections: A Dynamic General Equilibrium Approach. CEPR Discussion Paper No. 7202.

Diamond, D. 1984. "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies* 51 (3): 393–414.

Diamond, D., and P. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (3): 401–19.

Diamond, D., and R. Rajan. 2001. "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking." *Journal of Political Economy* 109 (2): 287–327.

Dib, A. 2009. "Banks, Credit Market Frictions, and Business Cycles." Manuscript, Bank of Canada (August).

European Central Bank. 2009. "Housing Finance in the Euro Area." ECB Occasional Paper No. 101.

Fagan, G., J. Henry, and R. Mestre. 2005. "An Area-wide Model for the Euro Area." *Economic Modelling* 22 (1): 39–59.

Francis, W., and M. Osborne. 2009. "Bank Regulation, Capital and Credit Supply: Measuring the Impact of Prudential Standards." Occasional Paper No. 36, United Kingdom Financial Services Authority.

Gambacorta, L., and P. E. Mistrulli. 2004. "Does Bank Capital Affect Lending Behavior?" *Journal of Financial Intermediation* 13 (4): 436–57.

Gerali, A., S. Neri, L. Sessa, and F. M. Signoretti. 2010. "Credit and Banking in a DSGE Model of the Euro Area." *Journal of Money, Credit, and Banking* 42 (s1): 107–41.

Gertler, M., and P. Karadi. 2009. "A Model of Unconventional Monetary Policy." Mimeo, New York University (April).

Gertler, M., and N. Kiyotaki. 2009. "Financial Intermediation and Credit Policy in Business Cycle Analysis." Mimeo, New York University (October).

Gordy, M., and B. Howells. 2006. "Procyclicality in Basel II: Can We Treat the Disease without Killing the Patient?" *Journal of Financial Intermediation* 15 (3): 395–417.

Gropp, R., and F. Heider. 2010. "The Determinants of Bank Capital Structure." *Review of Finance* 14 (4): 587–622.

Gropp, R., C. Kok Sørensen, and J. Lichtenberger. 2007. "The Dynamics of Bank Spreads and Financial Structure." ECB Working Paper No. 714.

Hülsewig, O., E. Mayer, and T. Wollmershäuser. 2009. "Bank Behavior, Incomplete Interest Rate Pass-Through, and the Cost Channel of Monetary Policy Transmission." *Economic Modelling* 26 (6): 1310–27.

Iacoviello, M. 2005. "House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle." *American Economic Review* 95 (3): 739–64.

Iacoviello, M., and S. Neri. 2010. "Housing Markets Spillovers: Evidence from an Estimated DSGE Model." *American Economic Journal: Macroeconomics* 2 (2): 125–64.

Jokivuolle, E., I. Kiema, and T. Vesala. 2009. "Credit Allocation, Capital Requirements and Procyclicality." Discussion Paper No. 23, Bank of Finland.

Kannan, P., P. Rabanal, and A. Scott. 2009. "Monetary and Macroprudential Rules in a Model with House Prices." IMF Working Paper No. 09/251.

Kashyap, A. N., and J. Stein. 2000. "What Do a Million Observations on Banks Say about the Transmission of Monetary Policy?" *American Economic Review* 90 (3): 407–28.

———. 2004. "Cyclical Implications of the Basel II Capital Standards." *Economic Perspectives* (Federal Reserve Bank of Chicago) 28 (1): 18–31.

Kishan, R. P., and T. P. Opiela. 2006. "Bank Capital and Loan Asymmetry in the Transmission of Monetary Policy." *Journal of Banking and Finance* 30 (1): 259–85.

Kiyotaki, N., and J. Moore. 1997. "Credit Cycles." *Journal of Political Economy* 105 (2): 211–48.

Kobayashi, K., T. Nakajima, and M. Inaba. 2007. "Collateral Constraint and News-Driven Cycles." Discussion Paper No. 07013, RIETI.

Kobayashi, T. 2008. "Incomplete Interest Rate Pass-Through and Optimal Monetary Policy." *International Journal of Central Banking* 4 (3): 77–118.

Kok Sørensen, C., and T. Werner. 2006. "Bank Interest Rate Pass-Through in the Euro Area: A Cross-Country Comparison." ECB Working Paper No. 580.

Liu, Z., P. Wang, and T. Zha. 2009. "Do Credit Constraints Amplify Macroeconomic Fluctuations?" Mimeo.

Lucas, R. E. 1990. "Liquidity and Interest Rates." *Journal of Economic Theory* 50 (2): 237–64.

Meh, C., and K. Moran. 2008. "The Role of Bank Capital in the Propagation of Shocks." Working Paper No. 2008-36, Bank of Canada.

Mester, L., and A. Saunders. 1995. "When Does the Prime Rate Change?" *Journal of Banking and Finance* 19 (5): 743–64.

Mojon, B. 2001. "Financial Structure and the Interest Rate Channel." *Economie et Prevision* 147 (1): 89–115.

Monacelli, T. 2009. "New Keynesian Models, Durable Goods and Collateral Constraints." *Journal of Monetary Economics* 56 (2): 242–54.

Notarpietro, A. 2007. "Credit Frictions and Household Debt in the US Business Cycle: A Bayesian Approach." Job Market Paper, Università Bocconi.

Peek, J., and E. S. Rosengren. 1995. "Bank Lending and the Transmission of Monetary Policy." In *Is Bank Lending Important for the Transmission of Monetary Policy?*, ed. J. Peek and E. S. Rosengren, 47–75. Federal Reserve Bank of Boston Conference Series No. 39.

Sander, H., and S. Kleimeier. 2006. "Expected versus Unexpected Monetary Policy Impulses and Interest Rate Pass-Through in Euro-Zone Retail Banking Markets." *Journal of Banking and Finance* 30 (7): 1839–70.

Santos, J. A. C., and A. Winton. 2009. "Bank Capital, Borrower Power, and Loan Rates." AFA 2010 Atlanta Meetings Paper. Available at http://papers.ssrn.com/so13/papers.cfm/abstractid=1343897.

Schmitt-Grohe, S., and M. Uribe. 2006. "Comparing Two Variants of Calvo-Type Wage Stickiness." NBER Working Paper No. 12740 (December).

Shi, S. 1997. "A Divisible Search Model of Fiat Money." *Econometrica* 65 (1): 75–102.

Smets, F., and R. Wouters. 2005. "Comparing Shocks and Frictions in US and Euro Area Business Cycles: A Bayesian DSGE Approach." *Journal of Applied Econometrics* 20 (2): 161–83.

———. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97 (3): 586–606.

Solomon, B.-D. 2011. "Firm Leverage, Household Leverage and the Business Cycle." Job Market Paper (February).

Stolz, S., and M. Wedow. 2005. "Banks' Regulatory Capital Buffer and the Business Cycle: Evidence for German Savings and Cooperative Banks." Discussion Paper (Banking and Financial Studies) No. 07/2005, Deutsche Bundesbank.

Van den Heuvel, S. J. 2002. "Does Bank Capital Matter for Monetary Transmission?" *Economic Policy Review* (Federal Reserve Bank of New York) 1 (1): 267–70.

———. 2008. "The Welfare Cost of Bank Capital Requirements." *Journal of Monetary Economics* 55 (2): 298–320.

Zhu, H. 2008. "Capital Regulation and Banks' Financial Decisions." *International Journal of Central Banking* 4 (1): 165–212.

Discussion of "Macroeconomic Propagation under Different Regulatory Regimes: Evidence from an Estimated DSGE Model of the Euro Area"

Andrew Powell
Research Department,
Inter-American Development Bank

It is obvious to any reader that the paper by Matthieu Darracq Pariès, Christoffer Kok Sørensen, and Diego Rodriguez-Palenzuela (this issue) represents an enormous amount of careful and detailed research. There is an old saying normally applied to theoretical papers in economics—that they are akin to a sausage, and that while the final product may be very good, one does not necessarily want to know what went into it. But perhaps the reverse is true for this type of paper. The research itself is perhaps as important as the results; what one gains from actually conducting this type of research may dominate what one learns from only considering the results. The other way to state this is that it is very difficult to convey everything that has been learnt from this type of research in the form of a standard written paper.

The paper develops a dynamic stochastic general equilibrium (DSGE) model with a large number of constituent elements. What are the main ingredients of the paper? The basic model has patient households that end up as savers and impatient ones that are the early consumers, and households may form habits; it has monopolistic unions and differentiated labor inputs; and it has entrepreneurs that produce residential goods (using capital, labor, and land) or non-residential goods (using capital and labor). There is a monopolistically competitive retail sector that sells to a perfectly competitive retail sector with a continuum of differentiated goods. The banking sector has three types of institutions, including a wholesale branch (assumed to be competitive), a (monopolistically competitive) deposit branch, and (competitive) lending branches.

The financial sector is the focus of much of the action in the paper. Loans are one period and, in different versions of the model, are at fixed or have state-contingent rates of interest. Loans are subject to strategic default, which in turn depends on a technology to seize the assets of the borrower. Loans to households are subject to a loan-to-value type restriction, and households may or may not be constrained in their borrowing decisions. Interest rates (as per prices) are set using Calvo-staggered contracts such that monetary policy has sluggish effects. Banks are subject to either Basel I or Basel II type minimum capital regulations. Finally, with respect to the government sector, the fiscal authorities finance expenditure with lump-sum taxes, and monetary policy is set using a type of Taylor rule.

If these are the main ingredients of the model, how is the dish then cooked? There are six types of uncertainty in the model: technology shocks (general, housing, non-housing, labor, and public expenditure), preference shocks (consumption and housing), price shocks (to price markups and bank spreads), idiosyncratic shocks (to households and entrepreneurs), bank capital shocks, and monetary policy shocks. The model has over 100 parameters. Many of these are fixed using the relevant literature to determine reasonable values, but some fifteen are estimated by calibrating the model to euro-area quarterly data.

The results of the paper are illustrated in a variety of ways including a set of impulse responses. The stated aim of the authors is to illustrate the role of financial frictions and also to consider the role of monetary policy and how monetary policy interacts with the different bank regulatory regimes (Basel I or Basel II and Basel III's anticyclical capital requirements). There are 20 impulse responses for each shock analyzed, and some 15 shocks are analyzed—so 300 impulse responses. In each graph there are 4 lines (baseline model, a version with high bank capital, another with imperfect interest rate pass-through, and finally a case with a predetermined interest rate), so some 1,200 lines to consider. In this rather brief discussion I will certainly not attempt to be comprehensive!

Indeed there is clearly a lot going on and much to consider and to discuss. There are so many ingredients mixed together that it reminds me of some comments Italian friends made to me regarding two Italian dishes—namely, polpette and pasticcio al forno. The former, polpette, is a tasty dish, and many ingredients can be thrown in, including the occasional leftover. It can be mixed up without too much care being taken, and it still tastes very good. Pasticcio al forno, however, is very different. It has many specific ingredients and they must be treated and added with extreme care and delicacy. Reading this paper, one wonders, is this polpette or is it pasticcio al forno? I will come back to this question below.

In fact, given the complexity of the model, the authors do a reasonably good job in providing some simple intuitions regarding the basic results of the paper. For example, a positive technology shock leads to a lower interest rate and greater activity through an accelerator-type effect. A negative shock to households increases default probability, increases interest rates, and dampens economic activity, although this is attenuated through monetary policy.

However, having said this, the paper is hard to follow at times. The art of DSGE modeling has taken different routes. A route favored by some academics has been to attempt to employ these tools to see what model features might explain particular features of the data. A good example is the paper by Aguiar and Gopinath (2007), who claim that introducing a stochastic growth term can capture the sort of economic fluctuations found particularly in emerging economies. One might speculate, given the recent financial crisis, that this argument could be more relevant to industrialized economies. An aim of this strand of the literature is to explain the data with a parsimonious model to help our understanding of what might be needed (i.e., the minimum required) to be able to characterize specific aspects of the data under consideration. Chang and Fernandez (2010), in an interesting recent paper, develop a model that encompasses Aguiar and Golpinath (2007), as it has the stochastic growth term but also a particular financial friction. They claim that the latter actually dominates in terms of more closely explaining the fluctuations in emerging economy data. This example shows how this strand of the literature is advancing by attempting to delineate improved model features that might capture aspects of the data more accurately.

A second route of DSGE modeling puts less emphasis on parsimony and more on explaining as many aspects of the data as possible. This is the route which tends to be favored by central banks that need to understand the wider impact of their actions on various parts of the economy. This paper is firmly within this strand of the literature. As already mentioned, the ingredients are both plentiful and complex.

A potential problem with this modeling choice is that one is then not sure which parts of the model are really doing the trick in terms of capturing different features of the data. For example, considering the impulse response graphs, I am struck by the apparent relatively small differences between the versions of the model with and without the collateral constraints. The authors allude to a Kiyotaki and Moore (1997) type credit cycle, where higher asset prices might relieve the constraint while a steep fall in asset prices would imply that the constraint binds and force banks to curtail lending with further negative consequences for the economy and asset prices. It would be of interest to know whether the apparent relatively small effects are simply due to the fact that this feature of the model is not really needed, given the EU 15 data used to calibrate the model, or to some other reason.

A second avenue worthy of further analysis is how the various financial frictions in the model interact. In an interesting recent paper, Martin and Taddie (2010) develop a model with two financial frictions and claim that their interaction amplifies the effects. Two frictions may then be worth more than their sum in terms of explaining economic fluctuations. The complexity of the model employed in this paper may not allow for such a clean theoretical analysis, but carefully chosen simulations could allow for exploration of this idea. As it stands, there is little in the paper on how the various model features interact.

The simulations on the effects of Basel II are of considerable interest. The conclusion is that Basel II (relative to Basel I) would have increased GDP volatility by some 5 percent, keeping monetary policy constant, although the authors also note a higher impact of shocks to the loan book due to the higher risk weights on riskier loans and banks will need to recapitalize more frequently. All in all, however, the effects of Basel II on volatility would have been quite small over the period.

An interesting finding is the interaction between macroprudential policies such as Basel III's anticyclical bank capital requirements and monetary policy. The authors find strong support for macroprudential policies (to minimize a loss function over growth, inflation, and interest rate volatility), and when macroprudential policies are operating, they conclude that monetary policy should not respond to asset prices or to credit. But they also state that the optimal rule would be difficult to implement in practice, as bank leverage would then become very volatile (4.8 times the baseline). In fact, bank leverage becomes quite volatile in the Basel II simulations (2.3 times the baseline) and GDP volatility becomes more volatile. Basel II then appears to introduce "bad" bank leverage volatility, whereas the macroprudential rule may produce "good" bank leverage volatility.

However, it is not entirely clear what is going on here. If macroprudential policy reduces economic fluctuations, one might expect bank leverage to become less volatile and not more. Banks should increase leverage less in the good times and reduce it less in the bad. This then begs the questions, what are banks actually doing in this model and is leverage volatility really a useful metric and a potential problem? The authors appear to be able to live with a Basel II simulation with leverage volatility being 2.3 times the baseline, but there is no convincing argument why 4.8 times the baseline is a problem. And if 2.3 times the baseline is okay, then why do the authors then feel the need to restrict leverage volatility all the way back to the baseline volatility for the case of macroprudential policy with restrictions on bank leverage changes?

I would also like to make a couple of more general points regarding DSGE modeling and macroprudential policies. In general, such models have actually found rather little impact for macroprudential policies including anticyclical capital regulations, and this paper in general reinforces this view. There may be at least two reasons for this. The first is that, on average, banks tend to hold buffers over actual capital requirements, and indeed those buffers are surely endogenous with respect to economic volatility. This is most clearly seen in the case of emerging markets. Take Latin America as an example. Broadly speaking, capital requirements around the continent are close to Brazil's 11 percent of assets at risk, substantially higher than Basel's recommended 8 percent minimum. But banks actually hold around 16 percent of capital in relation to assets at risk in the region.[1] Tier 1 ratios are also substantially higher than requirements and higher than in most G10 countries.

In this paper, banks target 11 percent of capital to assets at risk and are subject to a quadratic loss function that ensures that capital stays above the 8 percent minimum virtually always. In a recent paper I have written with coauthors (Aliaga-Diaz, Olivero, and Powell 2011), banks are forward looking and again are subject to a quadratic loss function if banks hit requirements and we calibrate this model to fit Latin American bank capital buffers. Repullo and Suarez (2008) develop a model taking a slightly different approach, as banks in their model hold buffers to ensure they can take advantage of profitable opportunities if they arise.

The fact that banks endogenously hold such buffers, which may then vary over lending cycles, dampens the effect of banks' procyclicality and will then also reduce the impact of anticyclical macroprudential policies. Hence calibrating these models to actual average bank capital buffers tends to lower the apparent value of macroprudential policies. One issue with this is that bank capital buffers vary across financial systems that tend to be as strong as their weakest link. To date, DSGE modeling has not captured well how weaker financial institutions and contagion, through financial contracts or investor reactions, may affect the behavior of all financial institutions and hence the real economy.

A second point relates to the interaction between the real economy, finance, and asset prices, which is critical to an understanding of credit cycles. To date, economic models (including DSGE models) have not captured the size of asset price movements, both up and down. In recent work, Cesa-Bianchi, Rebucci, and Powell (2011), using a global-VAR model, show that the majority of crises (stock market, currency, and banking)—including, for example, all U.S. stock market crises—are preceded by periods of "exuberance," defined as asset price booms that are not explained by economic fundamentals.

---

[1]See the International Monetary Fund's *Global Financial Stability Report* for data on bank capital as a percentage of assets at risk.

The point is that, to date, it is not clear that the current state of DSGE modeling really captures the types of asset price booms and busts that may be partly driving credit cycles. In turn, this may then mean that they do not find the critical role for macroprudential policies that many have in mind. It seems an odd comment regarding this paper, but perhaps there remains a missing ingredient.

To conclude, this is a very interesting paper with a great deal going on. I asked the question earlier whether it might be polpette or pasticcio al forno, and my answer is that in the end it is impossible to tell. This is in fact my main criticism of the paper as it stands. What I mean by this is that it is hard to tell which is the critical ingredient or ingredients and which aspects of the model are truly driving the results obtained. The paper, and the large amount of research conducted behind the paper, has significantly pushed forward the art of DSGE modeling into characterizing financial frictions and attempting to understand their impacts on the real economy. The paper includes many ingredients and in fact opens up several new avenues for future work. The authors may wish to focus on particular aspects of the model and simulations to attempt to understand more precisely what is actually driving specific results. This would then provide useful results to direct future work in the area, and I certainly look forward to reading future papers from the authors in this direction.

References

Aguiar, M., and G. Gopinath. 2007. "Emerging Market Business Cycles: The Cycle Is the Trend." *Journal of Political Economy* 115 (1): 69–102.

Aliaga-Diaz, R., M. Olivero, and A. Powell. 2011. "The Macroeconomic Effects of Anti-cyclical Bank Capital Requirements: Latin America as a Case Study." Mimeo, Inter-American Development Bank.

Cesa-Bianchi, A., A. Rebucci, and A. Powell. 2011. "Exuberance Crises." Mimeo, Research Department, Inter-American Development Bank.

Chang, R., and A. Fernandez. 2010. "On the Sources of Aggregate Fluctuations in Emerging Economies." NBER Working Paper No. 15938 (April).

Kiyotaki, N., and J. Moore. 1997. "Credit Cycles." *Journal of Political Economy* 105 (2): 211–48.

Martin, A., and F. Taddie. 2010. "International Capital Flows and Credit Market Imperfections: A Tale of Two Frictions." Mimeo. Available at http://www.hec.unil.ch/documents/seminars/deep/510.pdf.

Repullo, R., and J. Suarez. 2008. "The Procyclical Effects of Basel II." Mimeo, CEMFI.

Capital Regulation and Tail Risk*

Enrico Perotti[a], Lev Ratnovski[b], and Razvan Vlahu[c]
[a]University of Amsterdam, De Nederlandsche Bank, and Duisenberg School of Finance
[b]International Monetary Fund
[c]De Nederlandsche Bank

The paper studies risk mitigation associated with capital regulation, in a context where banks may choose tail risk assets. We show that this undermines the traditional result that higher capital reduces excess risk taking driven by limited liability. Moreover, higher capital may have an unintended effect of enabling banks to take more tail risk without the fear of breaching the minimal capital ratio in non-tail risky project realizations. The results are consistent with stylized facts about pre-crisis bank behavior, and suggest implications for the optimal design of capital regulation.

JEL Codes: G21, G28.

*We thank Douglas Gale (the editor), an anonymous referee, Morten Bech, Urs Birchler, Stijn Claessens, Giovanni Dell'Ariccia, Dale Gray, Alexander Guembel, Thomas Nitschka, Fausto Panunzi, Rune Stenbacka, Javier Suarez, Anjan Thakor, Andrew Winton, and Adam Zawadowski, as well as participants at seminars held at De Nederlandsche Bank, the International Monetary Fund, and at the 14th Swiss Society for Financial Markets conference, the 3rd IJCB Financial Stability conference, the 4th International Risk Management conference on New Dimensions in Risk Management, the 6th Finlawmetrics conference on Financial Regulation and Supervision, the IFABS conference on Financial Intermediation, Competition and Risk, the 26th meeting of the European Economic Association, the Sveriges Riksbank 2011 ESCB Day-Ahead conference on Financial Markets Research, the Federal Reserve Board and JMCB conference on Regulation of Systemic Risk, the 11th FDIC-JFSR Annual Bank Research Conference, the Bank of Finland-CEPR-JFI-SUERF conference on the Future of Risk Management, the CEPR-University of Vienna-Oesterreichische Nationalbank conference on Bank Supervision and Resolution, and the Bocconi University and JFI conference on Bank Competitiveness in the Post-Crisis World for helpful comments. Any errors are ours. The views expressed in this paper are those of the authors and should not be attributed to De Nederlandsche Bank or the International Monetary Fund. Vlahu gratefully acknowledges financial support by the Gieskes-Strijbis Foundation. Author e-mail addresses: e.c.perotti@uva.nl, lratnovski@imf.org, r.e.vlahu@dnb.nl.

1. Introduction

Regulatory reform in the wake of the recent financial crisis has focused on an increase in capital cushions of financial intermediaries. Basel III rules have doubled the minimal capital ratio and directed banks to hold excess capital as conservation and countercyclical buffers above the minimum (Bank for International Settlements 2010). These arrangements complement traditional moral suasion and individual targets used by regulators to ensure adequate capital cushions.

There are two key arguments in favor of higher capital. The first is an ex post argument: capital can be seen as a buffer that absorbs losses and hence reduces the risk of insolvency. This risk-absorption role also mitigates systemic risk factors, such as collective uncertainty over counterparty risk, which had a devastating propagation effect during the recent crisis. The second considers the ex ante effects of buffers: capital reduces limited-liability-driven incentives of bank shareholders to take excessive risk, by increasing their "skin in the game" (potential loss in case of bank failure; Holmstrom and Tirole 1997, Jensen and Meckling 1976).

Yet some recent experience calls for caution. First, banks are increasingly exposed to tail risk, which causes losses only rarely, but when those materialize they often exceed any plausible initial capital. Such risks can result from a number of strategies. A first example are carry trades reliant on short-term wholesale funding, which in 2007–08 produced highly correlated distressed sales (Gorton 2010). A second example is the reckless underwriting of contingent liabilities on systemic risk, callable at times of collective distress (Acharya and Richardson 2009). Finally, the combination of higher profits in normal times and massive losses occasionally arises in undiversified industry exposures to inflated housing markets (Shin 2009). A useful review of such strategies is provided in Acharya et al. (2009); International Monetary Fund (2010) highlights the importance of recognizing tail risk in financial stability analysis. Since under tail risk banks do not internalize losses independently of the level of initial capital, the buffer and incentive effects of capital diminish. Higher capital may become a less effective way of controlling bank risk.

Second, a number of major banks, particularly in the United States, appeared highly capitalized just a couple of years prior to the crisis. Yet these very intermediaries took excessive risks (often tail risk, or highly negatively skewed gambles). In fact, anecdotal evidence suggests that highly capitalized banks were looking for ways to put at risk their capital in order to produce returns for shareholders (Berger et al. 2008, Huang and Ratnovski 2009). Therefore, higher capital may create incentives for risk taking instead of mitigating them.

This paper seeks to study these concerns by reviewing the effectiveness of capital regulation and, in particular, of excess capital buffers (that is, above minimum ratios), in dealing with tail risk events. We reach two key results.

First, we show that the traditional buffer and incentives effects of capital become less powerful when banks have access to tail risk projects. The reason is that tail risk realizations can wipe out almost any level of capital. Left tails limit the effectiveness of capital as the absorbing buffer and restrict "skin in the game" because a part of the losses is never borne by shareholders. Hence, under tail risk, excess risk-shifting incentives of bank shareholders may exist almost independently of the level of initial or required capital.

Second, having established that under tail risk the benefits of higher capital are limited, we consider its possible unintended effects. We note that capital regulation also affects bank risk choices through the threat of capital adjustment costs when banks have to raise equity to comply with minimum capital ratios. (These costs are most commonly associated with equity dilution under asymmetric information on the value of illiquid bank assets—see Myers and Majluf 1984—or reduced managerial incentives for efficiency—Jensen 1986).[1] Similar to "skin in the game," capital adjustment costs make banks averse to risk and may discourage risky bank strategies. However, unlike "skin in the game," the incentive effects of capital adjustment costs fall with higher bank capital because the probability of breaching the minimal capital ratio decreases.

Of course, if highly capitalized banks internalized all losses, they would have taken risk only if that was socially optimal (would have offered a higher net present value, or NPV). Yet this result changes dramatically once we introduce tail risk. Then, even banks with

[1]The fact that adjustment costs of bank equity raising are significant was highlighted, for example, in the Basel Committee–FSF (2010) assessment of the impact of the transition to stronger capital requirements.

high capital never internalize all losses, and may take excess risk. Moreover, the relationship between capital and risk can become non-monotonic. The reason is interesting. In the first place, tail risk leads to insolvency whatever the initial bank capital, so higher capital does not sufficiently discourage risk taking for well-capitalized banks through "skin in the game." At the same time, higher excess capital allows banks to take the riskier projects without breaching the minimal capital ratio (and incurring large capital adjustment costs) in the case of low (non-tail) returns. So under tail risk, higher capital may create conditions where highly capitalized banks take more excess risk. Further, we show that the negative effect of extra capital on risk taking becomes stronger when banks get access to projects with even higher tail risk.

To close the model, we derive the bank's choice of initial capital in the presence of tail risk, and the implications for optimal capital regulation. We show that a bank may *choose* to hold higher capital in order to create a cushion over the minimal capital requirement so as to be able to take tail risk without the fear of a corrective action in case of marginally negative project realizations. Then, capital regulation has to implement two bounds on the values of bank capital: a bound from below (a minimal capital ratio) to prevent ordinary risk shifting and a bound from above (realistically, in the form of special attention devoted to banks with particularly high capital) in order to assure that they are not taking tail risk.

These results are interesting to consider in historic context. Most sources of tail risk that we described are related to recent financial innovations. In the past, tail risk in traditional loan-oriented depository banking was low (both project returns and withdrawals largely satisfied the law of large numbers), hence "skin-in-the-game" effects dominated, and extra capital led to lower risk taking. Yet now, when banks have access to tail risk projects, the buffer and "skin-in-the-game" effects that are the cornerstone of the traditional approach to capital regulation became weak, while effects where higher capital enables risk taking became stronger. Therefore, due to financial innovation, the beneficial effects of higher capital were reduced, while the scope for undesirable effects increased.

The paper has policy implications relevant for the current debate on strengthening capital regulation. The simpler conclusion is that it is impossible to control all aspects of risk taking using a single instrument. The problem of capital buffers is that they are effective as long as they can minimize not just the chance of default but also the loss given default. Contractual innovation in finance has enabled intermediaries to manufacture risk profiles which allow them to take maximum advantage of limited liability even with high levels of capital. The key to containing gambles with skewed returns is to either prohibit extreme bets or to increase their ex ante cost. Leading policy proposals now emerging are to charge prudential levies on strategies exposed to systemic risk (Acharya et al. 2010), such as extremely mismatched strategies (Perotti and Suarez 2009, 2010), or derivative positions written on highly correlated risks.

A more intricate conclusion relates to implications for capital regulation. The results do not imply that less capital is better: this was not the case in recent years. However, they suggest the following. First, regulators should acknowledge that traditional capital regulation has limitations in dealing with tail risk. This is similar, for example, to an already-accepted understanding that it has limitations in dealing with correlation risk (Acharya 2009). Second, banks with significant excess capital may be induced to take excess risk (in order to use or put at risk their capital), as amply demonstrated by the crisis experience. Hence, simply relying on higher and "excess" capital of banks as a means of crisis prevention may have ruinous effects if it produces a false sense of comfort. Finally, authorities should introduce complementary measures to target tail risk next to the policy on procyclical and conservation buffers. In this context, enhanced supervision with a focus on capturing tail risk may be essential.

We see our paper as related to two key strands of the banking literature. First are the papers on the unintended effects of bank capital regulation. Early papers (Kahane 1977, Kim and Santomero 1988, Koehn and Santomero 1980) took a portfolio optimization approach to banking and caution that higher capital requirements can lead to an increase in risk of the risky part of the bank's portfolio. Later studies focus on incentive effects. Boot and Greenbaum (1993) show that capital requirements can negatively affect asset quality due to a reduction in monitoring incentives. Blum (1999), Caminal and Matutes (2002), Flannery (1989), and Hellman, Murdock, and Stiglitz (2000) argue that higher capital can make banks take more risk as they attempt to compensate for the cost of capital.

Our paper follows this literature, with a distinct and contemporary focus on tail risk.[2] On the empirical front, Angora, Distinguin, and Rugemintwari (2009) and Bichsel and Blum (2004) find a positive correlation between levels of capital and bank risk taking.

The second strand consists of the recent papers on the regulatory implications of increased sophistication of financial intermediaries and the recent crisis. These papers generally argue that dealing with new risks (including systemic and tail risk) requires new regulatory tools (Acharya and Yorulmazer 2007, Acharya et al. 2010, Brunnermeier and Pedersen 2008, Huang and Ratnovski 2011, and Perotti and Suarez 2009, 2010).

The structure of the paper is as follows. Section 2 outlines the theoretical model. Section 3 describes the traditional "skin-in-the-game" effect of capital on risk taking. Section 4 shows how higher capital can enable risk taking when banks have access to tail risk projects. Section 5 endogenizes a bank's choice of initial capital and provides insights for optimal capital regulation. Section 6 concludes. The proofs and extensions are in the appendices.

2. The Model

The model has three main ingredients. First, the bank is managed by an owner-manager (hereafter, the banker) with limited liability, who can opportunistically engage in asset substitution. Second, the bank operates in a prudential framework based on a minimal capital ratio, with a capital adjustment cost if the bank fails to meet the ratio and has to raise extra equity. Finally, the bank has access to tail risk projects. Such a setup is a stylized representation of the key relevant features of the modern banking system. There are three dates $\left(0, \frac{1}{2}, 1\right)$, no discounting, and everyone is risk neutral.

The Bank. At date 0, the bank has capital C and deposits D. For convenience, we normalize $C + D = 1$. Deposits are fully insured

[2]Recent studies develop different measures for banks' tail risk. Acharya and Richardson (2009), Adrian and Brunnermeier (2009), and De Jonghe (2010) compute realized tail risk exposure over a certain period by using historical evidence of tail risk events, while Knaup and Wagner (2010) propose a forward-looking measure for bank tail risk.

at no cost to the bank; they carry a 0 interest rate and need to be repaid at date 1.[3]

The bank has access to two alternative investment projects. Both require an outlay of 1 at date 0 (all resources available to the bank) and produce return at date 1. The return of the *safe* project is certain: $R_S > 1$. The return of the *risky* project is probabilistic: high, $R_H > R_S$, with probability p; low, $0 < R_L < 1$, with probability $1-p-\mu$; or extremely low, $R_0 = 0$, with probability μ. We consider the risky project with three outcomes in order to capture both the second (variance) and the third (skewness, or "left tail," driven by the R_0 realization) moments of the project's payoff.

In the spirit of the asset substitution literature, we assume that the net present value (NPV) of the safe project is higher than that of the risky project:

$$R_S > pR_H + (1-p-\mu)R_L, \tag{1}$$

and yet the return on the safe asset, R_S, is not too high, so that the banker has incentives to choose the risky project at least for low levels of initial capital:

$$R_S - 1 < p(R_H - 1). \tag{2}$$

The left-hand side of (2) is the banker's expected payoff from investing in the safe project, and the right-hand side is the expected payoff from shifting to the risky project, conditional on the bank having no initial capital, $C = 0$ and $D = 1$. We consequently study conditions under which the bank's leverage creates incentives to opportunistically choose the suboptimal, risky project.

For definiteness, the bank chooses the safe project when indifferent. The bank has no continuation value beyond date 1. (We discuss the impact of a positive continuation value in appendix 2; it reduces bank risk taking but does not affect our qualitative results.)

The bank's project choice is unobservable and unverifiable. However, the return of the project chosen by the bank becomes observable and verifiable before final returns are realized,

[3]The presence of not fully risk-based deposit insurance is an inherent feature of most contemporary banking systems and one of the main rationales for bank regulation (Dewatripont and Tirole 1994).

at date $\frac{1}{2}$.[4] This allows the regulator to impose corrective action on an undercapitalized bank.

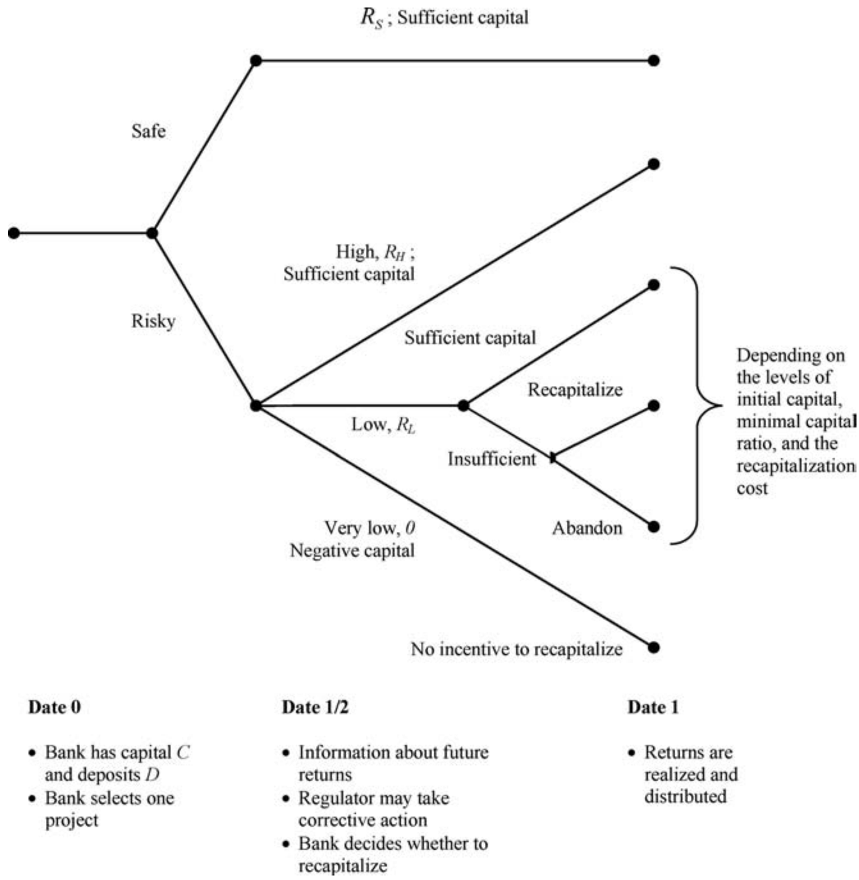
Capital Regulation. Capital regulation is based on the minimal capital (leverage) ratio. We take this regulatory design as exogenous, since it is the key feature of Basel regulation. We define the bank's capital ratio, $c = (A - D)/A$, where A is the value of bank assets, D is the face value of deposits, and $A - D$ is its economic capital. At date 0, before the investment is undertaken, the capital ratio is $c = C/(C + D) = C$. At dates $\frac{1}{2}$ and 1, the capital ratio is $c_i = (R_i - D)/R_i$, with $i \in \{S, H, L, 0\}$ reflecting project choice and realization. The fact that the date $\frac{1}{2}$ capital position is defined in a forward-looking way is consistent with the practice of banks recognizing known future gains or losses.

At any point in time, the bank's capital ratio c must exceed a minimum $c_{\min}$, $c_{\min} > 0$. We assume that the minimal ratio is satisfied at date 0: $c > c_{\min}$. Consequently, the minimal ratio is also satisfied for realizations R_S (when the bank chooses the safe project) and R_H (when the bank chooses the risky project and is successful): $c_H > c_S > c > c_{\min}$, since $R_H > R_S > 1$. The minimal capital ratio is never satisfied for R_0 (in the extreme low outcome of the risky project), since the bank's capital is negative, $c_0 = -\infty < 0 < c_{\min}$. In a low realization of the risky project R_L the bank's capital sufficiency is ambiguous. As we will show below, depending on the bank's initial capital, it can be positive and sufficient, $c_L > c_{\min}$; positive but insufficient, $0 < c_L < c_{\min}$; or negative, $c_L < 0$.

The regulator imposes corrective action at date $\frac{1}{2}$ if a bank fails to satisfy the minimal capital ratio. The banker is given two options. One is to surrender the bank to the regulator. Then, the bank's equity value is wiped out and the banker receives a zero payoff. Alternatively, the bank can attract additional capital to bring its capital ratio to the regulatory minimum, $c_{\min}$. We assume that attracting capital carries a cost for the existing bank shareholder. The cost reflects the dilution when equity issues are viewed by new investors as negative signals, or when there is a downward-sloping demand for the bank's shares. Both factors may be especially strong when

[4]The assumption that project choice is unobservable while project returns are is a standard approach to modeling (Hellman, Murdock, and Stiglitz 2000, Rochet 2004).

Figure 1. The Timeline

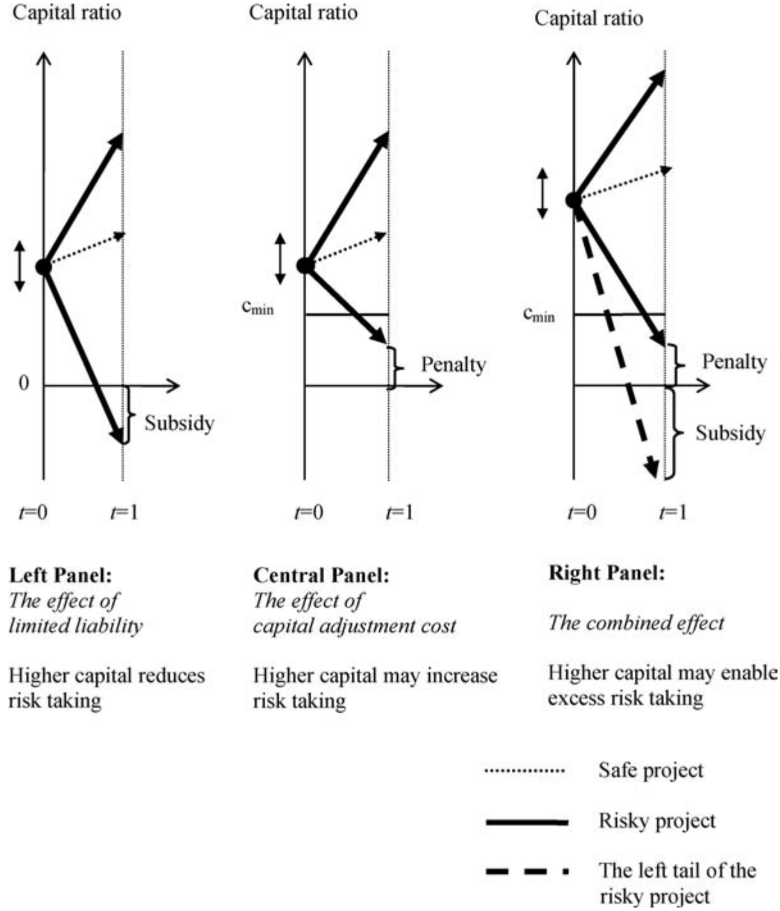


the offering is performed with urgency. The presence of such costs is well established in the literature (Asquith and Mullins 1986). In the main model, we treat the cost of recapitalization as fixed at T. In appendix 2 we discuss a specification with concave cost (i.e., the cost of recapitalization falls with higher bank capital) and show that it does not affect our results. The banker chooses to abandon rather than recapitalize the bank when indifferent.

Timeline. The model outcomes and the sequence of events are depicted in figure 1.

Intuition. Figure 2 provides a simple, illustrative intuition for the effects that we intend to capture in our formal analysis. Consider

Figure 2. The Two Opposite Effects of Capital on Bank Risk Taking



a bank that chooses between a safe and a risky project, and note how the bank's level of initial capital affects that choice. The classic Myers and Majluf (1984) channel focuses on the consequences of limited liability, which subsidizes risk taking and tilts the bank's incentives towards a risky project. Then, higher capital discourages risk taking by making shareholders internalize more of the bank's losses in the bad outcome ("skin in the game," the left panel). We introduce an additional effect associated with the minimal capital requirements. A bank with positive but insufficient capital is subject to costly corrective action: shareholders have to recapitalize or abandon the bank. This penalizes risky projects. Then, a bank with higher capital may choose *more* risk, because higher capital reduces the probability of breaching the minimal capital ratio (the capital adjustment cost effect, the center panel). Of course, if a highly capitalized bank internalized all the costs of risk taking, it would choose the risky project only if that was socially optimal (offered a higher NPV). To formalize the excess risk taking of highly capitalized banks, we combine the two effects in a framework where a bank's risky project can both marginally breach the minimal capital ratio (triggering a capital adjustment cost) and result in an extremely negative outcome (tail risk, which falls under the limited-liability constraint, the right panel). We find that, as a result of the combination of the two effects, the relationship between bank capital and risk taking may become non-linear. In particular—the key result that we will emphasize—banks with higher capital may choose inefficiently high risk when such risk has a significant tail component.

The return function with three outcomes is the simplest form that supports the insights of this model. The distinction between the marginal bad (low) and the extreme bad (tail) realization is necessary to simultaneously capture the effects of aversion to recapitalization and risk shifting. Our results can also arise in more general distributions, including continuous, risky return distributions having similar features: a mass in the left tail and a possibility of marginally negative outcomes.

3. "Skin in the Game" and Tail Risk

In this section we show that the traditional "skin-in-the-game" incentive effects of higher capital on risk taking become weaker when the bank has access to tail risk projects. This brings us to the first policy result, that capital regulation may have limited effectiveness in dealing with tail risk. Throughout the section, we abstract from the effects of capital adjustment costs (we assume no minimum capital ratio), which we introduce in the next section. We solve the model backwards, first deriving the payoffs depending on bank project choice and then the project choice itself. The solution is followed by comparative statics and a calibration exercise.

3.1 Payoff and Project Choice

Consider the bank with access to a tail risk project ($\mu \geq 0$), in a setup with no capital adjustment costs ($T = 0$). The banker's payoff from choosing the safe project is

$$\Pi_S^{T=0} = R_S - D = R_S - (1 - c). \tag{3}$$

The banker's payoff from choosing the risky project is

$$\Pi_R^{T=0} = p \cdot [R_H - (1 - c)] + (1 - p - \mu) \cdot \max\{R_L - (1 - c); 0\}, \tag{4}$$

where on the right-hand side of (4) the first term is the expected payoff in R_H realization, and the second term is the expected payoff in R_L realization. The third realization, R_0, occurs with probability μ and carries a zero payoff.

The bank chooses the safe project over the risky project when

$$\Pi_S^{T=0} \geq \Pi_R^{T=0},$$

which is equivalent to

$$R_S - (1 - c) \geq p \cdot [R_H - (1 - c)] + (1 - p - \mu) \cdot \max\{[R_L - (1 - c)]; 0\}. \tag{5}$$

The following proposition describes the bank's investment decision.

PROPOSITION 1. *The bank's project choice depends on its initial capital c:*

(a) For

$$R_S < pR_H + (1 - p)R_L, \tag{6}$$

the bank chooses the safe project if

$$c \geq 1 - \frac{R_S - pR_H - (1 - p - \mu)R_L}{\mu}$$

and the risky project otherwise.

(b) For

$$R_S \geq pR_H + (1-p)R_L, \tag{7}$$

the bank chooses the safe project if

$$c \geq 1 - \frac{R_S - pR_H}{1-p}$$

and the risky project otherwise.

Proof. See appendix 1.

The intuition for case (b) of the above proposition is that when R_S is high enough, the bank's risk-shifting incentives are so low that the bank will only take a risky project when it has negative capital under the R_L realization, allowing it to shift more of the downside to the creditors. Then, the bank gets the same zero payoff in the R_0 and R_L realizations, and its project choice is not affected by the tail risk probability μ. Case (b) therefore represents the case of negligible tail risk. We therefore further focus on case (a), which allows us to study the impact of tail risk on the bank's project choice. We denote

$$c^{T=0} = 1 - \frac{R_S - pR_H - (1-p-\mu)R_L}{\mu}, \tag{8}$$

with $c^{T=0}$ being the threshold for risk-shifting incentives under (6).

3.2 Comparative Statics

We study how the threshold $c^{T=0}$, the initial capital necessary to prevent the bank from risk shifting, is affected by the project's tail risk μ. To maintain comparability, we consider transformations of the risky project that increase μ but preserve its expected value, denoted by $E(R)$. There are various ways to alter model parameters to achieve that, but we highlight the two with the best interpretations, which we analyze in turn.

CASE 1. *Some of the sources of tail risk in the recent crisis were carry trades or undiversified exposures to housing markets. Such activities transform the distribution of the risky project towards extreme outcomes: within the confines of our model, we can interpret that as a shift in the probability mass from R_L to R_0 and R_H.*

Formally, that implies an increase in μ and p, at the expense of $(1 - p - \mu)$. To keep $E(R)$ constant, following an increase in μ by $\Delta\mu$, p should increase by $\frac{R_L}{R_H - R_L}\Delta\mu$.

Using (8), we find that

$$\left.\frac{\partial c^{T=0}}{\partial \mu}\right|_{E(R)=constant} = \frac{R_S - E(R)}{\mu^2} > 0. \tag{9}$$

So the amount of capital necessary to prevent risk shifting increases in tail risk.

CASE 2. *Another source of tail risk was the underwriting of contingent liabilities on market risk; in this case the bank obtains ex ante premia (higher return) in all cases when the tail risk is not realized. Formally, this can be interpreted as a higher μ compensated by higher R_L and R_H, so that $E(R)$ is constant. In order to achieve this, following an increase in μ by $\Delta\mu$, both R_L and R_H should increase by $R_L \frac{\Delta\mu}{1-\mu-\Delta\mu}$.*

Similarly to the previous case, using (8), we find that

$$\left.\frac{\partial c^{T=0}}{\partial \mu}\right|_{E(R)=constant} = \frac{R_S - E(R)}{\mu^2} > 0.$$

Hence, again, the amount of capital necessary to prevent risk shifting increases in tail risk.

In both cases, observe that $c^{T=0}$ grows logarithmically in μ.[5] Therefore, capital becomes progressively a less effective incentive tool for controlling bank risk taking as tail risk μ increases, with the effect most pronounced for low values of μ. As an implication, tail risk limits the effectiveness of capital regulation in dealing with bank risk-taking incentives.

[5] We rewrite $c^{T=0}(\mu) = 1 - \frac{const}{\mu}$, with $const = R_S - E(R)$. The degree of polynomial $c^{T=0}(\mu)$ is given by $\lim_{\mu\to\infty} \frac{\mu \cdot [c^{T=0}(\mu)]'}{c^{T=0}(\mu)}$. This equals 0, the degree of the logarithm function.

3.3 Economic Significance: A Quantitative Example

The comparative statics exercise verified that as banks are able to take projects with higher tail risk μ, the buffer and incentive effects of capital diminish. Thus, in order to prevent banks from taking tail risk projects using minimal capital-based ("skin-in-the-game") incentives only, banks will need progressively higher levels of initial capital $c^{T=0}$. This section attempts to highlight the economic significance of these results through a simple calibration exercise.

Consider the following calibration parameters: $R_S = 1.03$, $R_H = 1.14$, $R_L = 0.92$, $p = 50$ percent, and $\mu = 1$ percent. Then, the expected return on the safe project is 3 percent, the expected return on the risky project is 2.1 percent, and the minimal level of capital necessary to prevent risk shifting is $c^{T=0} = 8$ percent. We take these parameter values as representing the case of low (or usual) tail risk.

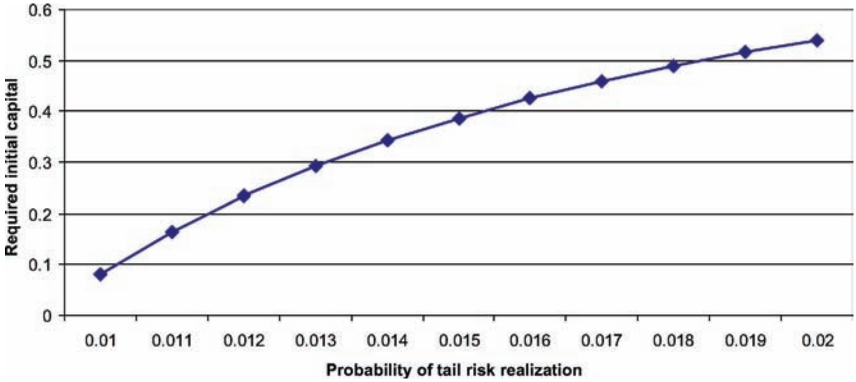
We ask how $c^{T=0}$ changes if tail risk μ increases, holding the expected value of the risky project fixed, as in the comparative statics exercise, by adjusting p to compensate for higher μ (case 1 in section 3.2). The result of the calibration exercise is shown in figure 3. As μ increases, so does $c^{T=0}$, and the increase in $c^{T=0}$ is economically significant. Indeed, even an increase in μ from 1 percent to 1.1 percent increases the initial capital necessary to prevent risk shifting from 8 percent to 16.4 percent. A doubling of μ to 2 percent increases the necessary initial capital to as much as 54 percent. Such values of initial capital are likely implausible in practice.

The calibration confirmed that, at least under some plausible circumstances, minimal capital requirements alone cannot prevent banks from taking excess tail risk, because the level of capital necessary for that would need to be implausibly high. In the next section, we study how the costly corrective action on undercapitalized banks may complement the capital requirements in dealing with bank risk taking.

4. Tail Risk and the Unintended Effects of Higher Capital

In the previous section, we showed that capital becomes a less effective tool for controlling bank risk taking in the presence of tail risk. We will now introduce an additional feature—capital adjustment

Figure 3. Tail Risk and the Initial Capital Required to Prevent Risk Shifting



costs—to obtain a stronger result. In addition to being a less powerful tool, higher capital may have unintended effects of *enabling* banks' risk taking. Specifically, we show that marginally capitalized banks do not take risk, because they are averse to breaching the minimal capital ratio in mildly negative realizations of the risky project (R_L). Yet banks with higher capital can take more risk, because their chance of breaching the ratio in such realizations is lower. Further, in comparative statics, we demonstrate that the unintended effects of higher capital are stronger when banks get access to projects with higher tail risk.

As before, we solve the model backwards: first we derive the payoffs depending on bank project choice, then the project choice itself. The solution is followed by comparative statics.

4.1 Payoffs and the Recapitalization Decision

The banker's payoff from choosing the safe project is

$$\Pi_S = R_S - D = R_S - (1 - c).$$

Now consider the banker's payoff from the risky project. When the project returns R_H, the banker obtains $R_H - (1 - c)$. When the project returns R_0, the banker obtains zero.

The case when the risky project returns R_L is more complex because, depending on the relative values of c and R_L, the bank's capital may be positive and sufficient, positive but insufficient, or negative. Consider these in turn.

Under R_L, the bank has positive and sufficient capital ($c_L \geq c_{\min}$) when

$$\frac{R_L - (1-c)}{R_L} \geq c_{\min},$$

which gives

$$c \geq c^{Sufficient} = 1 - (1 - c_{\min})R_L. \quad (10)$$

Then, the bank continues to date 1, repays depositors, and obtains $R_L - (1 - c)$.

When $c < c^{Sufficient}$, R_L leaves the bank with insufficient capital, $c_L < c_{\min}$, so it has to be abandoned or recapitalized at cost T. The banker chooses to recapitalize the bank for

$$R_L - (1 - c) - T > 0, \quad (11)$$

where the left-hand side is the banker's return after repaying depositors net off the recapitalization cost, and the right-hand side is the zero return in case the bank is abandoned. Expression (11) can be rewritten as

$$c > c^{Recapitalize} = 1 + T - R_L. \quad (12)$$

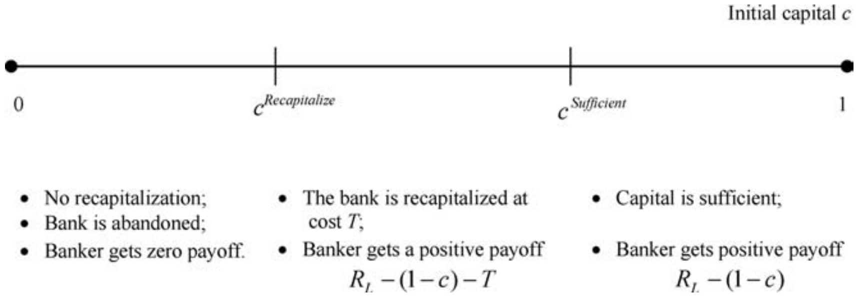
We focus our analysis on the case when $c^{Recapitalize} < c^{Sufficient}$, corresponding to

$$T < c_{\min}R_L, \quad (13)$$

so that there exist values of c: $c^{Recapitalize} < c < c^{Sufficient}$, where the banker chooses to recapitalize the bank in the R_L realization, instead of abandoning it. When T is larger than $c_{\min}R_L$, the banker always abandons a bank with insufficient capital. Note that both thresholds $c^{Recapitalize}$ and $c^{Sufficient}$ are in the $(0, 1)$ interval.

Figure 4 illustrates the bank's recapitalization decision.

Figure 4. Bank's Recapitalization Decision and Payoffs



Notes: The bank's recapitalization decision and banker's payoffs as a function of initial capital c, upon the realization of low return R_L, are as follows. For $c \geq c^{Sufficient}$, the bank has positive and sufficient capital at date 1/2. The bank continues to date 1, repays depositors, and obtains a positive payoff. For $c < c^{Sufficient}$, the bank has positive and insufficient or negative capital. The bank can be either abandoned or recapitalized. The bank is abandoned for $c \leq c^{Recapitalize}$. As a result, the bank is closed and the banker gets a zero payoff. The bank is recapitalized at a cost for $c^{Recapitalize} < c < c^{Sufficient}$. The bank continues to date 1, repays depositors, pays the recapitalization cost, and obtains a positive payoff.

Overall, the banker's payoff in the R_L realization of the risky project is

$$\Pi_L = \begin{cases} R_L - (1-c), \text{ if } c \geq c^{Sufficient} \\ R_L - (1-c) - T, \text{ if } c^{Recapitalize} < c < c^{Sufficient} \\ 0, \text{ if } c \leq c^{Recapitalize} \end{cases}, \quad (14)$$

and the overall payoff of the risky project is

$$\Pi_R = p \cdot [R_H - (1-c)] + (1-p-\mu) \cdot \Pi_L. \quad (15)$$

4.2 Project Choice

We now consider the bank's project choice at date 0, depending on its initial capital c. The bank chooses the safe project over the risky one for

$$\Pi_S \geq \Pi_R,$$

which is equivalent to

$$R_S - (1-c) \geq p \cdot [R_H - (1-c)] + (1-p-\mu) \cdot \Pi_L. \quad (16)$$

To describe the results, we introduce two thresholds:

$$W = pR_H + (1-p)R_L - \mu c_{\min} R_L \tag{17}$$

and

$$Z = pR_H + (1-p)(R_L - T) + \mu(T - c_{\min} R_L). \tag{18}$$

W is a threshold point for the presence of risk shifting in bank with high capital. For $R_S < W$ there exist values of initial capital such that even a well-capitalized bank with $c \geq c^{Sufficient}$ may still engage in risk shifting. Z is a threshold point for the presence of the capital adjustment cost effect. For $R_S \geq Z$ there exist values of initial capital such that a less capitalized bank ($c^{Recapitalize} < c < c^{Sufficient}$) may choose a safe project to prevent recapitalization costs upon the R_L realization of the risky project. The derivation of the thresholds is in appendix 1; the appendix also verifies that $Z < W$.

Then, the risk-shifting and capital adjustment effect of bank project choice interact with each other as follows:

PROPOSITION 2. *The bank's project choice is characterized by thresholds c^* and c^{**}:*

(a) For $Z \leq R_S < W$, there exist $c^* < c^{Sufficient}$ and $c^{**} > c^{Sufficient}$, such that

- *for $c < c^*$ the bank chooses the risky project and may abandon or recapitalize it upon the R_L realization;*
- *for $c^* \leq c < c^{Sufficient}$ the bank chooses the safe project to avoid abandonment or recapitalization upon the R_L realization; the choice of the safe project here represents the capital adjustment cost effect;*
- *for $c^{Sufficient} \leq c < c^{**}$ the bank chooses the risky project because its capital is high enough to avoid breaching the minimal capital ratio in the R_L realization; the choice of the risky project here represents risk shifting enabled by higher capital; and*
- *for $c \geq c^{**}$ the bank chooses the safe project because its capital is high enough to prevent risk shifting.*

(b) For $R_S < Z$, there exists $c^{**} > c^{Sufficient}$ such that for $c < c^{**}$ the bank chooses the risky project and for $c \geq c^{**}$

the safe project. There is only a risk-shifting effect: a bank with $c < c^{Sufficient}$ *never chooses a safe project to avoid recapitalization cost.*

(c) *For* $R_S \geq W$*, there exists* $c^* < c^{Sufficient}$ *such that for initial capital* $c < c^*$ *the bank chooses a risky project and for* $c \geq c^*$ *the safe project. There is only a capital adjustment cost effect: a bank with* $c > c^{Sufficient}$ *never engages in risk shifting.*

Proof. See appendix 1.

The thresholds,

$$c^* = 1 - \frac{R_S - pR_H - (1 - p - \mu)(R_L - T)}{\mu} \tag{19}$$

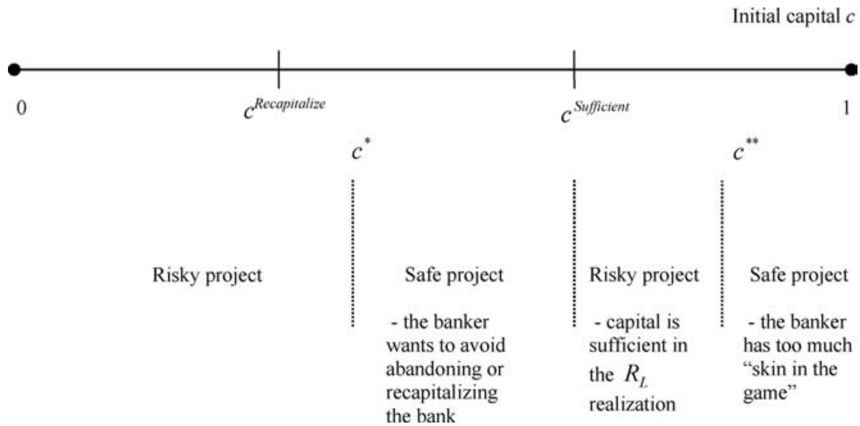
and

$$c^{**} = 1 - \frac{R_S - pR_H - (1 - p - \mu)R_L}{\mu}, \tag{20}$$

are also derived in appendix 1.

Case (a) of proposition 2 contains the main result of our paper: that the relationship between bank capital and risk taking can be non-monotonic in the presence of tail risk and capital adjustment cost. When capital is very low, $c < c^*$, the banker faces strong risk-shifting incentives and a low cost of abandoning the bank, and hence chooses high risk. For intermediate initial capital, $c^* \leq c < c^{Sufficient}$, the banker's equity value is higher, and the banker chooses a safe project to avoid abandoning or recapitalizing the bank in the R_L realization. The choice of the safe project is driven by capital adjustment cost—a novel effect highlighted in this paper. Yet as soon as the bank has initial capital high enough to satisfy the minimal ratio in the R_L realization, for $c^{Sufficient} \leq c < c^{**}$, the capital adjustment cost stops being binding and the banker again switches to the risky project, driven by the risk-shifting effect. Finally, for very high levels of capital, $c \geq c^{**}$, the banker has so much skin in the game that risk-shifting incentives are not binding. This is the traditional effect of capital regulation; recall that under tail risk, c^{**} may be prohibitively very high. The bank's project choice is depicted in figure 5.

Figure 5. Bank's Project Choice



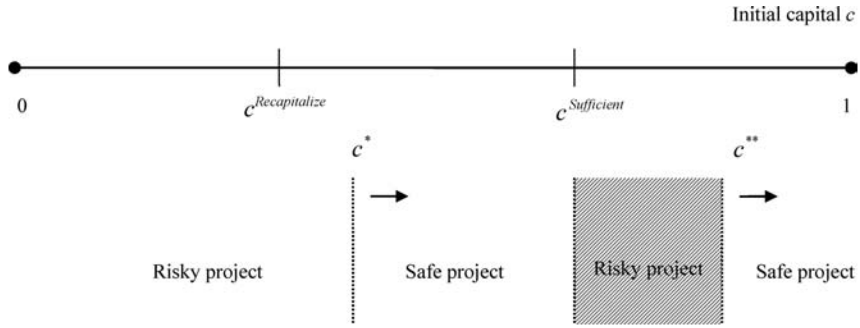
Notes: The bank's project choice depending on the level of initial capital, in case (a) of proposition 2: The relationship between bank capital and risk taking is non-linear and is characterized by two thresholds as follows. When the level of capital is low ($c < c^*$), the bank prefers the risky project, while for high level of capital ($c \geq c^{**}$), the safe project is chosen. For intermediate level of capital ($c^* \leq c < c^{**}$), the bank prefers either the safe project (for $c^* \leq c < c^{Sufficient}$) or the risky one (for $c^{Sufficient} \leq c < c^{**}$).

4.3 Comparative Statics

In this section we repeat the comparative statics exercise of section 3.2, in the presence of capital adjustment costs—with respect to case (a) of proposition 2. We show that when tail risk increases (the risky project has a heavier left tail), highly capitalized banks get stronger incentives to take excess risk. We use the two transformations of the risky project highlighted in section 3.2.

CASE 1. *When a higher μ is compensated by a higher p, keeping $E(R)$ constant, that affects both thresholds c^* and c^{**}. To focus on bankers' incentives to take excessive risk, we consider the interval $[c^{Sufficient}, c^{**})$, corresponding to levels of initial bank capital for which the bank undertakes the risky project. Note that $c^{Sufficient}$ is determined only by $c_{\min}$ and R_L (see (10)), and hence is unaffected by a change in the probability distribution of the risky project. The critical threshold for the discussion is therefore c^{**}. From (8),*

Figure 6. Bank's Project Choice when the Risky Project Has a Heavier Left Tail: Case 1



Notes: A heavier left tail is characterized by a higher probability for the extremely low outcome (i.e., a higher μ). A change in the return profile of the risky project following a change in probability distribution (i.e., both p and μ are increased, all else equal, such that the expected value of the risky project remains the same) affects both thresholds c^* and c^{**}. The interval $[c^{Sufficient}, c^{**})$ widens, suggesting that well-capitalized banks which behave prudently in absence of tail risk projects have a strong incentive to undertake more risk, if projects with a heavier left tail are available in economy.

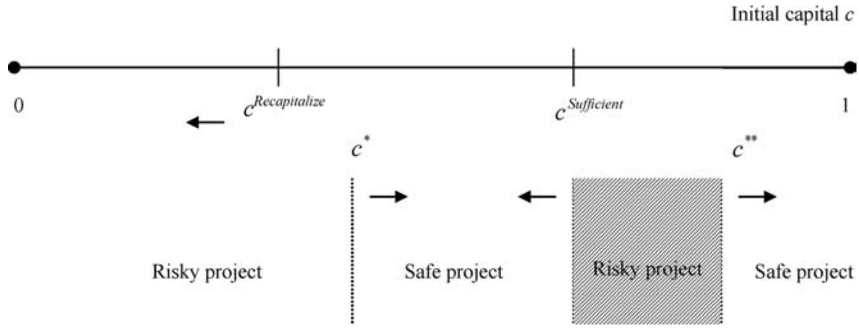
$c^{**} = c^{T=0}$. *The impact of the change in probability distribution of the risky project on* c^{**} *is the same as in (9):*

$$\frac{\partial c^{**}}{\partial \mu}\bigg|_{E(R)=constant} = \frac{R_S - E(R)}{\mu^2} > 0.$$

This means that when tail risk increases, the interval $[c^{Sufficient}, c^{**})$ on which a well-capitalized bank chooses the risky project expands. Interestingly, the interval expands because banks with *higher* capital start taking more risk. This highlights the relationship between tail risk and the unintended effects of higher bank capital. The intuition is that when investment returns become more polarized, they enable well-capitalized banks to earn higher profits in good time, while at the same time reducing the expected cost of recapitalization since the intermediate low return R_L is less frequent. Unintended effects of bank capital affect specifically the well-capitalized banks. Figure 6 illustrates the case.

CASE 2. *When a higher* μ *is compensated by higher* R_L *and* R_H*, this change in the return profile of the risky asset affects*

Figure 7. Bank's Project Choice when the Risky Project Has a Heavier Left Tail: Case 2



Notes: A heavier left tail is characterized by a higher probability for the extremely low outcome (i.e., a higher μ). A change in the return profile of the risky project following a change in probability distribution (i.e., μ is increased), compensated by higher R_L and R_H, all else equal, such that the expected value of the risky project remains the same, affects thresholds c^* and c^{**}, and $c^{Sufficient}$ as well. The interval $[c^{Sufficient}, c^{**})$ widens even more, suggesting that both well- and less-capitalized banks will start choosing the risky project.

thresholds c^*, c^{**}, and $c^{Sufficient}$. To focus on the incentives of well-capitalized banks to take excessive risk, we again consider the interval $[c^{Sufficient}, c^{**})$. Note that $c^{Sufficient}$ is decreasing in R_L (see (10)) and hence in μ. At the same time, from (9), c^{**} is increasing in μ. Hence here the interval $[c^{Sufficient}, c^{**})$ widens even more in μ than in case 1, and both more- and less-capitalized banks start choosing the risky projects. Figure 7 illustrates the case.

5. Bank Capital Choices and Optimal Capital Regulation

We have previously identified how a bank's incentives to take tail risk depend on its initial level of capital. We can now study how the ability of banks to take tail risk affects bank capital choices and what the implications are for the optimal capital regulation. To endogenize bank capital choice, we need to introduce the bank's cost of holding capital. We assume the following:

(A) The bank's private cost of holding capital is $c\gamma$, $\gamma > 1$. This cost represents the alternative cost of using the banker's own funds elsewhere; see Hellmann, Murdock, and Stiglitz (2000). The assumption assures that, all else equal, the bank will want to hold as little capital as possible.

(B) The cost of bank capital becomes prohibitive for high values of capital, making $c_{\min} = c^{**}$ impossible to implement. Recall that under tail risk, c^{**} (the level of capital necessary to prevent bank risk taking solely through the "skin-in-the-game" channel) can become very high (section 3.3).

We can now formulate the result on the bank's endogenous choice of initial capital. We focus on case (a) of proposition 2, where the relationship between bank capital and risk taking is non-monotonic in bank capital.

LEMMA 1. *Setting $c_{\min} < c^*$ is never optimal (because the bank will always choose $c_{\min}$ and a risky project); therefore $c_{\min} \geq c^*$.*

Proof. It follows from proposition 2 and assumption A.

We can now formulate the result on the bank's capital choice:

PROPOSITION 3. *For $c_{\min} \geq c^*$, the bank's private capital choice is either the minimal capital $c_{\min}$ or the level of capital sufficient to avoid recapitalization costs in the R_L realization of the risky project $c^{Sufficient}$. There exists $\gamma^* > 1$:*

$$\gamma^* = 1 + \mu \frac{R_L}{1 - R_L} - \frac{R_S - pR_H - (1 - p - \mu)R_L}{(1 - c_{\min})(1 - R_L)} \tag{21}$$

such that the bank chooses $c^{Sufficient}$ for $\gamma < \gamma^*$ and $c_{\min}$ otherwise.

Proof. See appendix 1.

The intuition is as follows. Because capital is costly, the bank will choose the lowest capital possible on each of the intervals $[c_{\min}, c^{Sufficient})$ and $[c^{Sufficient}, c^{**})$. Capital below $c_{\min}$ is ruled out by capital regulation under lemma 1, and capital at or above c^{**} is ruled out by assumption B above. To establish the bank's choice for capital, we therefore have to compare profits at two points: $c_{\min}$ and $c^{Sufficient}$. The banker will prefer $c^{Sufficient}$ if the cost of maintaining extra capital, proportional to γ, is not too high.

It is important to understand the economic logic behind the choice between $c_{\min}$ and $c^{Sufficient}$. The point $c_{\min}$ gives lower capital, so the bank saves on its cost. But the bank has to take the safe project (socially optimal but less beneficial to shareholders) to avoid the capital adjustment cost. Should a bank switch to $c^{Sufficient}$, it would voluntarily choose to incur the cost of holding higher capital, because that would *enable the bank to take higher risk.* Indeed, the bank will choose the risky project because it is no longer constrained by the threat of capital adjustment cost. We have therefore established that in the presence of tail risk, banks may choose to accumulate capital *in order* to be able to take tail risk.

With this in mind, we can now turn to optimal capital regulation. Recall that the level of capital that allows us to rule out bank risk taking through "skin in the game" only, c^{**}, is not plausible. Therefore, the only objective of regulation that is feasible in our setup is to assure that the bank's capital is in the $[c^*, c^{Sufficient})$ interval, where a bank takes the safe project due to its aversion to capital adjustment costs. Both below and above this interval, the bank will undertake the risky project. Such a regulatory outcome can be implemented with two instruments. The first is a standard minimal capital requirement, set at $c_{\min} = c^*$ (there is no reason to set $c_{\min} > c^*$, since capital is costly). The second is an effective constraint on the bank's excess capital over the minimal capital requirement. The constraint can be interpreted, in practice, not as a limit, but as special attention that regulators should give to banks with excess capital, anticipating that such banks are more likely than others to take tail risk. We summarize with the following conjecture:

CONJECTURE 1. *In the presence of tail risk, optimal capital regulation combines a minimal capital requirement $c_{\min} = c^*$ and effective constraints on banks with excess capital (above $c^{Sufficient}$) to prevent them from taking tail risk.*

6. Conclusion

This paper examined the relationship between bank capital and risk taking when banks have access to tail risk projects. We showed that traditional capital regulation becomes less effective in controlling bank risk because banks never internalize the negative realizations of tail risk projects. Moreover, we have suggested novel channels for unintended effects of higher capital: it enables banks to take higher tail risk without the fear of breaching the minimal capital requirement in mildly bad (i.e., non-tail) project realizations. The results are consistent with stylized facts about pre-crisis bank behavior and have implications for the design of bank regulation.

Appendix 1. Proofs

Proof of Proposition 1

Consider first the case when $c > 1 - R_L$. The relevant incentive compatibility condition derived from (5) becomes $R_S - (1-c) \geq p\cdot[R_H - (1-c)] + (1-p-\mu)\cdot[R_L - (1-c)]$, which can be rewritten as $c \geq 1 - \frac{R_S - pR_H - (1-p-\mu)R_L}{\mu}$. We denote $c^{T=0} = 1 - \frac{R_S - pR_H - (1-p-\mu)R_L}{\mu}$. The threshold exists and is strictly larger than $1-R_L$ for values of R_S satisfying (6). Hence, for $R_S < pR_H + (1-p)R_L$, $\exists\, c^{T=0} \in (1-R_L, 1]$ such that $\forall\, c \in (1-R_L, c^{T=0})$ the risky project is selected and $\forall\, c \in [c^{T=0}, 1]$ the safe project is chosen. Otherwise, if (6) is not fulfilled, the bank selects the safe project $\forall\, c \in (1-R_L, 1]$.

Consider now the case $c \leq 1 - R_L$. The relevant incentive compatibility condition derived from (5) is $R_S - (1-c) \geq p\cdot[R_H - (1-c)]$. The condition is equivalent with $c \geq 1 - \frac{R_S - pR_H}{1-p}$. We denote $c^{T=0}_{Traditional} = 1 - \frac{R_S - pR_H}{1-p}$. The threshold exists and is below or equal to $1 - R_L$ for values of R_S satisfying (7). Thus, for $R_S \geq pR_H + (1-p)R_L$, $\exists\, c^{T=0}_{Traditional} \in [0, 1-R_L]$ such that $\forall\, c \in [0, c^{T=0}_{Traditional})$ the risky project is selected and $\forall\, c \in [c^{T=0}_{Traditional}, 1-R_L]$ the safe project is chosen. Otherwise, if (7) is not fulfilled, the bank selects the risky project $\forall\, c \in [0, 1-R_L]$.

To sum up, when (6) is fulfilled, $\exists\, c^{T=0} \in (1-R_L, 1]$ such that $\forall\, c \in [0, c^{T=0})$ the risky project is selected and $\forall\, c \in [c^{T=0}, 1]$ the safe project is chosen. Likewise, when (6) is not fulfilled, $\exists\, c^{T=0}_{Traditional} \in [0, 1-R_L]$ such that $\forall\, c \in [0, c^{T=0}_{Traditional})$ the risky project is selected and $\forall\, c \in [c^{T=0}_{Traditional}, 1]$ the safe project is chosen.

Proof of Proposition 2

We consider two scenarios in turn. We start by analyzing a scenario in which the cost of recapitalization is such that

$\frac{\mu}{1-p}c_{\min}R_L < T < c_{\min}R_L$. Subsequently we show that our results are similar for $T \leq \frac{\mu}{1-p}c_{\min}R_L$. Both cases follow from assumption (13) of low adjustment cost.

The Case of $\dfrac{\mu}{1-p}c_{\min}R_L < T < c_{\min}R_L$

We study a bank's behavior for three levels of initial capital: low (i.e., $c \leq c^{Recapitalize}$), intermediate (i.e., $c^{Recapitalize} < c < c^{Sufficient}$), and high (i.e., $c \geq c^{Sufficient}$).

Consider first the case when $c \in [0, c^{Recapitalize}]$. The banker never finds it optimal to recapitalize for low realization of the risky project. The relevant incentive compatibility condition derived from (16) is $R_S - (1-c) \geq p \cdot [R_H - (1-c)]$, where the left-hand side is the return on investing in the safe project and the right-hand side is the expected return on selecting the risky project. The condition can be rewritten as $c \geq 1 - \frac{R_S - pR_H}{1-p}$. We denote $c_1^* = 1 - \frac{R_S - pR_H}{1-p}$. The threshold c_1^* exists if and only if the next two constraints are jointly satisfied:

$$R_S < 1 - p + pR_H, \tag{22}$$

$$R_S > pR_H + (1-p)(R_L - T). \tag{23}$$

The former condition guarantees a positive c_1^*, while the latter forces the threshold to be lower than $c^{Recapitalize}$, the upper limit for the interval we analyze. If (22) is not fulfilled, then $c_1^* < 0$ and $\forall\ c \in [0, c^{Recapitalize}]$ the bank prefers the safe project. If (23) is not fulfilled, then $c_1^* > c^{Recapitalize}$ and $\forall\ c \in [0, c^{Recapitalize}]$ the bank invests risky. When both constraints are simultaneously satisfied, $\exists\ c_1^* \in [0, c^{Recapitalize}]$ such that $\forall\ c \in [0, c_1^*)$ the risky project is selected and $\forall\ c \in [c_1^*, c^{Recapitalize}]$ the safe project is chosen. Assumption (2) implies that (22) is always fulfilled.

Consider now the case when $c \in (c^{Recapitalize}, c^{Sufficient})$. The banker finds it optimal to recapitalize for low realization R_L. The relevant incentive compatibility condition is $R_S - (1-c) \geq p \cdot [R_H - (1-c)] + (1-p-\mu) \cdot [R_L - (1-c) - T]$, with the certain return on choosing the safe project depicted on the left-hand side and expected return on investing in the risky project depicted on the right-hand side. Rearranging terms, the condition can be rewritten as $c \geq 1 - \frac{R_S - pR_H - (1-p-\mu)(R_L - T)}{\mu}$. We denote $c_2^* = 1 - \frac{R_S - pR_H - (1-p-\mu)(R_L - T)}{\mu}$. Similarly with the previous case, the threshold c_2^* exists if and only if it is simultaneously higher and lower than the lower and the higher boundary of the analyzed interval, respectively. The conditions are as follows:

$$R_S < pR_H + (1-p)(R_L - T), \tag{24}$$

$$R_S > pR_H + (1-p)(R_L - T) + \mu(T - c_{\min} R_L). \tag{25}$$

Condition (24) is the opposite of (23). Thus, a satisfied condition (23) implies that (24) is not fulfilled. Condition (24) not satisfied implies that $c_2^* < c^{Recapitalize}$ and $\forall\ c \in (c^{Recapitalize}, c^{Sufficient})$ the bank prefers the safe project. If the second condition is not fulfilled, then $c_2^* > c^{Sufficient}$ and $\forall\ c \in (c^{Recapitalize}, c^{Sufficient})$ the bank invests risky. When both constraints are simultaneously satisfied, $\exists\ c_2^* \in (c^{Recapitalize}, c^{Sufficient})$ such that $\forall\ c \in (c^{Recapitalize}, c_2^*)$ the risky project is selected. The safe project is preferred $\forall\ c \in [c_2^*, c^{Sufficient})$.

Consider now the final interval, when $c \in [c^{Sufficient}, 1]$. For low realization R_L, the bank always complies with the regulatory requirements. No additional capital is needed. The relevant incentive compatibility condition is $R_S - (1-c) \geq p \cdot [R_H - (1-c)] + (1 - p - \mu) \cdot [R_L - (1-c)]$. Rearranging terms, the condition becomes $c \geq 1 - \frac{R_S - pR_H - (1-p-\mu)R_L}{\mu}$. We denote $c^{**} = 1 - \frac{R_S - pR_H - (1-p-\mu)R_L}{\mu}$. The threshold c^{**} exists if and only if $c^{**} > c^{Sufficient}$ and $c^{**} < 1$. The latter is always fulfilled following from the assumption (1) of higher NPV for the safe project. The former condition is depicted in (26) below. When (26) is not satisfied, the bank prefers the safe project for any level of initial capital larger than $c^{Sufficient}$. Otherwise, $\forall\ c \in [c^{Sufficient}, c^{**})$ the risky project is selected, while the safe project is preferred $\forall\ c \in [c^{**}, 1]$.

$$R_S < pR_H + (1-p)R_L - \mu c_{\min} R_L. \tag{26}$$

Next we discuss the process of project selection. Recall that $Z = pR_H + (1-p)(R_L - T) + \mu(T - c_{\min} R_L)$ and $W = pR_H + (1 - p)R_L - \mu c_{\min} R_L$, from (18) and (17), respectively. We also denote $B = pR_H + (1-p)(R_L - T)$. Under assumption (13), $Z < B < W$. We distinguish among four possible scenarios:

- Scenario S1: $R_S < Z$. As a consequence, condition (25) is not satisfied and $\forall\ c \in (c^{Recapitalize}, c^{Sufficient})$ the bank selects the risky project. $R_S < Z$ also implies that $R_S < B$ and $R_S < W$. Condition (23) is not satisfied but (26) is. As a result, the bank invests risky $\forall\ c \in [0, c^{Recapitalize}] \cup [c^{Sufficient}, c^{**})$, and the bank invests safe $\forall\ c \in [c^{**}, 1]$.
- Scenario S2: $Z \leq R_S \leq B$. The right-hand side implies that condition (23) is not satisfied. For initial capital c lower than $c^{Recapitalize}$, the bank prefers the risky project. The left-hand side implies that condition (25) is fulfilled. Also, condition (24) is satisfied, being the opposite of (23). Hence, we can conclude that $\exists\ c^* \in (c^{Recapitalize}, c^{Sufficient})$ with $c^* = c_2^*$, such that $\forall\ c \in (c^{Recapitalize}, c^*)$ the risky project is selected, while the safe project is preferred $\forall\ c \in [c^*, c^{Sufficient})$. Condition (26) is also satisfied. Similarly with the previous scenario, the bank invests risky $\forall\ c \in [c^{Sufficient}, c^{**})$ and safe $\forall\ c \in [c^{**}, 1]$.
- Scenario S3: $B < R_S < W$. The left-hand side implies that condition (23) is satisfied. We can argue that $\exists\ c^* \in (0, c^{Recapitalize})$ with $c^* = c_1^*$, such that $\forall\ c \in [0, c^*)$ the risky project is selected, while the safe project is preferred $\forall\ c \in [c^*, c^{Recapitalize}]$. Condition (23) implies that (24) is not satisfied. Thus, $\forall\ c \in (c^{Recapitalize}, c^{Sufficient})$ the safe project will be selected. The bank investment decision is identical to the one from previous scenarios when the level of capital is high enough (i.e., c larger than $c^{Sufficient}$).
- Scenario S4: $R_S \geq W$. Neither condition (26) nor condition (24) are satisfied anymore. The bank selects the safe project $\forall\ c \in (c^{Recapitalize}, 1]$. However, condition (23) is fulfilled. Hence, $\exists\ c^* \in [0, c^{Recapitalize}]$ with $c^* = c_1^*$, such that $\forall\ c \in [0, c^*)$ the risky project is selected, while the safe project is preferred $\forall\ c \in [c^*, c^{Recapitalize}]$.

The values for thresholds c^* and c^{**} for case (a) of proposition 2 are derived under scenario S2 above for $Z \leq R_S \leq B$.

The Case of $T \leq \dfrac{\mu}{1-p} c_{\min} R_L$

We consider now the scenario under which the cost of recapitalization is very low. Lowering T has no quantitative impact on $c^{Sufficient}$

and $c^{Recapitalize}$, the thresholds in initial capital which trigger a bank's decision between raising additional capital or letting the regulator overtake the bank. Their relative position is unchanged: $c^{Sufficient}$ is larger than $c^{Recapitalize}$, following from easily verifiable identity $\frac{\mu}{1-p}c_{\min}R_L < c_{\min}R_L$ combined with our restriction on T. However, the process of project selection under assumption (13) is marginally affected. In this scenario $Z < W < B$, as a consequence of lower T. As discussed before, we distinguish among four possible scenarios: (S1') $R_S < Z$, (S2') $Z \leq R_S < W$, (S3') $W \leq R_S < B$, and (S4') $R_S > B$. Discussions for scenarios S1', S2', and S4' are identical to our previous discussion for scenarios S1, S2, and S4. We discuss scenario S3' next. $W \leq R_S$ implies that condition (26) is not satisfied. Hence, the bank prefers the safe project for any level of initial capital larger than $c^{Sufficient}$. $R_S < B$ implies that condition (23) is not satisfied. For initial capital c lower than $c^{Recapitalize}$, the bank prefers the risky project. However, condition (24) is satisfied, being the opposite of (23), and also condition (25) is implied by the fact that $W > Z$. Hence, we can conclude that $\exists\ c^* \in (c^{Recapitalize}, c^{Sufficient})$ with $c^* = c_2^*$, such that $\forall\ c \in (c^{Recapitalize}, c^*)$ the risky project is selected, while the safe project is preferred $\forall\ c \in [c^*, c^{Sufficient})$.

Analysis of the Robustness of Results of Proposition 2

We offer here a discussion for the results of proposition 2. We analyze a bank's project choice for the case of high cost of recapitalization: $T > c_{\min}R_L$; we show that our results are robust to this specification. Recall that under assumption (13), there exist values of c such that $c^{Recapitalize} < c < c^{Sufficient}$, where the banker chooses to recapitalize the bank following the R_L realization instead of abandoning it. We show next that when T is larger than $c_{\min}R_L$, the banker always abandons a bank with insufficient capital. Although the main results from proposition 2 are not qualitatively affected, higher recapitalization cost has a quantitative impact on our results. Therefore, we start by deriving the new conditions which drive these results. It is optimal for the bank to raise additional capital (if this was demanded by the regulator) when conditions $c_L < c_{\min}$ and (11) are simultaneously satisfied. The former condition implies that

$c < 1 - R_L(1 - c_{\min})$, while from the latter $c > 1 + T - R_L$. Under our modified assumption of high cost of recapitalization T, these conditions cannot be satisfied simultaneously. For any levels of initial capital c below $1 - R_L(1 - c_{\min})$, the banker receives a request for adding extra capital but never finds it optimal to do so because such an action will not generate positive payoffs. As a result, the bank is closed and the shareholder expropriated. Conversely, when the level of initial capital is above $1 - R_L(1 - c_{\min})$, the banking authority doesn't take any corrective action against the bank since returns R_L are above the critical level $R_{\min}$. We denote

$$c_{NEW}^{Recapitalize} = 1 - R_L(1 - c_{\min}), \tag{27}$$

where $c_{NEW}^{Recapitalize} \in (0, 1)$. Next, we explore the bank's project choice for levels of initial capital below and above this critical threshold.

Consider first the case when $c \in (0, c_{NEW}^{Recapitalize})$. The bank never recapitalizes for the low realization of the risky project. The bank would have incentive to select the safe project when $R_S - (1 - c) \geq p[R_H - (1 - c)]$, which implies that initial capital c to be larger than $1 - \frac{R_S - pR_H}{1-p}$. We previously denoted $c_1^* = 1 - \frac{R_S - pR_H}{1-p}$. This threshold exists if and only if (22) and the following condition are jointly satisfied:

$$R_S > pR_H + (1 - p)R_L(1 - c_{\min}). \tag{28}$$

The second condition guarantees that c_1^* is lower than $c_{NEW}^{Recapitalize}$, the upper boundary for the interval we analyze. For large returns on the safe project (i.e., condition (22) is not fulfilled), $\forall\, c \in (0, c_{NEW}^{Recapitalize})$ the bank prefers the safe project. If (23) is not fulfilled, then $\forall\, c \in (0, c_{NEW}^{Recapitalize})$ the bank invests risky. Otherwise, when both constraints are simultaneously satisfied, $\forall\, c \in (0, c_1^*)$ the risky project is selected and $\forall\, c \in (c_1^*, c_{NEW}^{Recapitalize})$ the safe project is chosen. Our assumption (2) implies that (22) is always fulfilled.

Consider now the second case when $c \in (c_{NEW}^{Recapitalize}, 1)$. The bank always complies with the regulatory requirements when R_L is obtained due to high initial capital. No additional capital is needed. The bank would have incentive to select the safe project when $R_S - (1 - c) \geq p[R_H - (1 - c)] + (1 - p - \mu)[R_L - (1 - c)]$, which implies $c \geq 1 - \frac{R_S - pR_H - (1-p-\mu)R_L}{\mu}$. We previously denoted $c^{**} = 1 - \frac{R_S - pR_H - (1-p-\mu)R_L}{\mu}$. The threshold c^{**} exists if and only if condition (26) is satisfied. The safe project is preferred for any level of initial capital larger than $c_{NEW}^{Recapitalize}$ whenever (26) is not satisfied. Otherwise, $\forall\ c \in (c_{NEW}^{Recapitalize}, c^{**})$ the risky project is selected, while the safe project is preferred $\forall\ c \in (c^{**}, 1)$.

Recall that $W = pR_H + (1-p)R_L - \mu c_{\min} R_L$. We also denote $Q = pR_H + (1-p)R_L(1 - c_{\min})$. It is easy to show that $Q < W$ due to the identity $1 - p - \mu > 0$. We distinguish among only three possible scenarios:

- Scenario S1": $R_S \leq Q$. As a consequence, condition (28) is not satisfied and $\forall\ c \in (0, c_{NEW}^{Recapitalize})$ the bank selects the risky project. $R_S < Q$ implies that $R_S < W$. Condition (26) is satisfied. As a result, the bank invests risky $\forall\ c \in (c_{NEW}^{Recapitalize}, c^{**})$ while preferring the safe project $\forall\ c \in (c^{**}, 1)$.
- Scenario S2": $Q < R_S < W$. The left-hand side implies that condition (28) is satisfied. This implies that $\exists\ c^* \in (0, c_{NEW}^{Recapitalize})$ with $c^* = c_1^*$, such that $\forall\ c \in (0, c^*)$ the risky project is selected, while the safe project is preferred $\forall\ c \in (c^*, c_{NEW}^{Recapitalize})$. Similarly with the previous scenario, the bank invests risky $\forall\ c \in (c_{NEW}^{Recapitalize}, c^{**})$, and safe $\forall\ c \in (c^{**}, 1)$. This result is implied by R_S being lower than W.
- Scenario S3": $R_S \geq W$. Condition (28) is satisfied, while condition (26) is not. Hence, the bank selects the risky projects $\forall\ c \in (0, c^*)$, with $c^* = c_1^*$, and the safe project $\forall\ c \in (c^*, 1)$.

To conclude, we can argue that the qualitative results of proposition 2 are valid under the relaxed assumption. Nevertheless, condition (25) has to be replaced by the relevant condition (28).

Bank's Choice when the Return on the Safe Project Is Large

Let us consider here that the return on the safe asset is large—that is, $R_S > 1 - p + pR_H$. This drives the following results under assumption (13): (i) condition (22) is not satisfied, implying that $\forall\ c \in [0, c^{Recapitalize}]$ the bank prefers the safe project; (ii) condition (23) is satisfied, which implies that condition (24) is not, and as a result $\forall\ c \in (c^{Recapitalize}, c^{Sufficient})$ the bank prefers the safe project; (iii) condition (26) is not satisfied and as a consequence $\forall\ c \in [c^{Sufficient}, 1]$ the bank invests in the safe project. Summing up, for any levels of initial capital c, the bank prefers the safe project when the certain return R_S is high enough.

Proof of Proposition 3

From equation (10), $c^{Sufficient} > c_{\min}$. We consider three cases for the level of $c_{\min}$. Consider first the case when $c_{\min} \leq c^{Recapitalize}$. From case (a) of proposition 2, the banker finds it optimal to select the risky project when the level of initial capital is in this region. The banker's expected payoff is $p \cdot [R_H - (1-c)] - c\gamma$, which is decreasing in initial capital c (the first derivative is negative, since $\gamma > 1$ by assumption (A)). Hence, under minimal capital ratio regulation, the banker chooses $c_{\min}$ as initial capital.

Consider now the case when $c_{\min} \in (c^{Recapitalize}, c^*)$. Again, from case (a) of proposition 2, the banker finds it optimal to select the risky project and recapitalize in the R_L realization. The banker's expected payoff is $p \cdot [R_H - (1-c)] + (1-p-\mu)[R_L - (1-c) - T] - c\gamma$, which is decreasing in initial capital c. Hence, under minimal capital ratio regulation, the banker chooses again $c_{\min}$ as initial capital.

Consider now the last case when $c_{\min} \in [c^*, c^{Sufficient})$. The banker finds it optimal to take no risk. The safe project is selected and the expected payoff for the banker is $R_S - (1-c) - c\gamma$. The expected payoff decreases in c. However, the banker can be better off selecting a higher level of initial capital. Consider that the banker decides to hold $c^{Sufficient}$. This allows risk taking (see case (a) from proposition 2). The expected payoff is $p \cdot [R_H - (1 - c^{Sufficient})] + (1-p-\mu) \cdot [R_L - (1 - c^{Sufficient})] - c^{Sufficient}\gamma$. The banker is better off selecting a higher level of capital if and only if

$$p \cdot [R_H - (1 - c^{Sufficient})] + (1 - p - \mu) \cdot [R_L - (1 - c^{Sufficient})] - c^{Sufficient}\gamma > R_S - (1 - c_{\min}) - c_{\min}\gamma. \quad (29)$$

Rearranging terms in (29), the condition becomes $\gamma < 1 + \mu\frac{R_L}{1-R_L} - \frac{R_S - pR_H - (1-p-\mu)R_L}{(1-c_{\min})(1-R_L)}$. We denote $\gamma^* = 1 + \mu\frac{R_L}{1-R_L} - \frac{R_S - pR_H - (1-p-\mu)R_L}{(1-c_{\min})(1-R_L)}$. The threshold γ^* is higher than 1 for $R_S < W$, with W given in (17)—see also case (a) of proposition 2 for further details. Therefore, we can conclude that $\exists\ \gamma^* \in (1, \infty)$ such that $\forall\ \gamma \in (1, \gamma^*)$ the banker selects $c = c^{Sufficient}$ and $\forall\ \gamma \in [\gamma^*, \infty)$ the banker selects $c = c_{\min}$.

Appendix 2. Extensions

We offer here two extensions for our model and examine the implications of charter value and of different specification for recapitalization costs. We show that our results are robust to these generalizations.

Charter Value

In section 2 we have assumed that there is no charter value for the continuation of bank's activity. In this section we introduce a positive charter value $V > 0$ and show that our results are robust to this extension. Our model suggests that low competition in banking, which provides a high charter value, leads to investment in the efficient safe project even by well-capitalized banks.

The role of banks' franchise values have been shown relevant in other studies. Hellmann, Murdock, and Stiglitz (2000) and Repullo (2004) argue that prudent behavior can be facilitated by increasing banks' charter value. They study the links between capital requirements, competition for deposits, charter value, and risk-taking incentives, and they point out that banks are more likely to gamble and to take more risk in a competitive banking system, since competition erodes profits and implicitly the franchise value. A similar idea is put forward by Matutes and Vives (2000). They argue that capital regulation should be complemented by deposit rate regulation and direct asset restrictions in order to efficiently keep risk taking under control. Acharya (2003) explores how continuation value affects risk preferences in the context of optimal regulation, and demonstrates the disciplining effect of charter value on bank risk taking. Finally, Furlong and Kwan (2005) and Keeley (1990) explore empirically the relation between charter value and different measures of bank risk, and they find strong evidence that bank charter value disciplined bank risk taking.

In the new setting, the banker's payoff from the safe project after repaying depositors becomes $\Pi_S^V = R_S - (1-c) + V$. The banker's payoff from the risky project is as follows: when R_H is realized, the banker gets $\Pi_H^V = R_H - (1-c) + V$, while the payoff is 0 for extremely low realization R_0. When the low return R_L is realized and capital is positive but insufficient ex post, the banker prefers to recapitalize at a cost T for lower levels of initial capital c. The reason for this is that the banker's expected payoff increases by V if the bank is not closed by the regulator. Hence, the bank raises additional capital when initial capital c is higher than $c_V^{Recapitalize}$, where

$$c_V^{Recapitalize} = 1 + T - R_L - V, \tag{30}$$

and $c_V^{Recapitalize} < c^{Recapitalize}$. On the other hand, the threshold point $c^{Sufficient}$ does not change, since it is given by the exogenous regulation.

We make the simplifying assumption that the charter value is not larger than a certain threshold:

$$V < 1 + T - R_L. \tag{31}$$

This makes threshold $c_V^{Recapitalize}$ positive and assures the existence for the area $[0, c_V^{Recapitalize}]$ where the bank is abandoned for low realization of the risky project. Consider the area $(c_V^{Recapitalize}, c^{Sufficient})$. When initial capital c is in this range, a bank that is subject to a regulator's corrective action prefers to raise additional capital. Since $c_V^{Recapitalize} < c^{Recapitalize}$, while the right boundary of the interval is left unchanged by any increase in V, we can argue that any reduction in banking competition, which increases bank charter value, makes the decision to raise fresh capital more likely.

We introduce the following two thresholds:

$$Z_V = pR_H + (1-p)(R_L - T) + \mu(T - V - c_{\min} R_L), \tag{32}$$

as the new threshold for the binding impact of the prompt corrective action (with $Z_V < Z$), and

$$B = pR_H + (1-p)(R_L - T). \tag{33}$$

Following a similar proof as for proposition 2, we can show that there exist two thresholds c_V^* and c_V^{**} for the level of initial bank capital such that under assumption (13) and for levels of return on the safe project satisfying $Z_V < R_S < B$, with Z_v and B defined in (32) and (33), respectively, the bank's investment preference is as follows:

(a) The bank prefers the risky project for $0 \leq c < c_V^*$, while for $c_V^* \leq c < c^{Sufficient}$ the safe project is preferred, with $c_V^* \in (c_V^{Recapitalize}, c^{Sufficient})$, where $c_V^{Recapitalize}$ and $c^{Sufficient}$ are defined in (24) and (10), respectively, and

$$c_V^* = 1 - V - \frac{R_S - pR_H - (1 - p - \mu)(R_L - T)}{\mu}. \qquad (34)$$

(b) The bank prefers the risky project for $c^{Sufficient} \leq c < c_V^{**}$ and the safe project for $c \geq c_V^{**}$, where $c_V^{**} \in (c^{Sufficient}, 1)$ and

$$c_V^{**} = 1 - V - \frac{R_S - pR_H - (1 - p - \mu)R_L}{\mu}. \qquad (35)$$

Observe that a positive charter value has a negative impact on all relevant thresholds which drive banks' preferences (i.e., $c_V^{Recapitalize}$, c_V^*, and c_V^{**}), except for $c^{Sufficient}$. Hence, we can argue that higher charter value plays the role of a counterbalancing force to the risk-taking incentives generated by the presence of risky projects with heavy left tails. This means that when the continuation value of banks' activity is high enough, both intervals $(0, c_V^*)$ and $(c^{Sufficient}, c_V^{**})$ shrink. This suggests that low competition in the banking industry induces banks with larger capital buffers to take less risk.

In summary, the results of our basic model are therefore robust to the introduction of charter value, conditional on the fact that this value is not too large. Large values of charter value reduce risk-taking incentives even for well-capitalized banks.

Concave Capital Adjustment Cost

In section 2 we considered a simple fixed cost of recapitalization. We now show that results are robust to a more general specification of this cost function.

In this section we discuss a variation of the model in which the cost of recapitalization has a fixed and a variable component. The variable component is proportional to the amount of new capital that the bank has to raise in order to comply with the minimal capital ratio. Specifically, the bank has to raise a capital level $R_{\min}-R_L$, where $R_{\min}$ equals

$$R_{\min} = \frac{1-c}{1-c_{\min}}. \tag{36}$$

The above threshold is derived from the condition of a minimal capital ratio of $c_{\min}$ (i.e., $c \leq c_{\min} = [R_{\min} - (1-c)]/R_{\min}$), by solving for the value of bank's assets (i.e., $R_{\min}$).[6] In this new setting, the recapitalization cost is concave in capital level and has the following specification:

$$Cost(c, R_L) = T + \beta\left(\frac{1-c}{1-c_{\min}} - R_L\right). \tag{37}$$

We assume that variable cost of recapitalization (i.e., β) is positive and not so low as to make the banker abandon the bank regardless the level of initial capital:

$$T < R_L(1+\beta). \tag{38}$$

The banker's payoff from the safe project, as well as the realizations of the risky project, are the same as in the basic model. However, when the low realization R_L is obtained, the bank is abandoned more often than in the basic model due to higher cost of recapitalization. The bank is closed when $c < c_{CC}^{Recapitalize}$, where

$$c_{CC}^{Recapitalize} = 1 + \frac{T - R_L(1+\beta)}{1 + \frac{\beta}{1-c_{\min}}}, \tag{39}$$

(with CC for concave cost).

[6]Consider the following example. Assume that the bank has to raise δ units of capital to satisfy the regulatory minimum when R_L is realized. Hence, $c_{\min} = \frac{R_L-(1-c)+\delta}{R_L+\delta}$. This implies that $R_L + \delta = \frac{1-c}{1-c_{\min}}$, which equals $R_{\min}$ according to (36). We can conclude that $\delta = R_{\min} - R_L$.

Under assumption (13), $c_{CC}^{Recapitalize} > c^{Recapitalize}$. On the other hand, the level of capital which guarantees that the bank satisfies ex post the regulatory minimal upon realization of R_L (i.e., $c^{Sufficient}$) remains unchanged. Hence, the interval $(c_{CC}^{Recapitalize}, c^{Sufficient})$ shrinks, suggesting that the bank is less likely to raise additional capital if required to do so.

We denote

$$B_{CC} = pR_H + (1-p)\frac{R_L(1+\beta) - T}{1 + \frac{\beta}{1-c_{\min}}}. \quad (40)$$

Following the lines of proof for proposition 2, we can show that there exist two thresholds c_{CC}^{*} and c_{CC}^{**} for the level of initial bank capital such that under assumption (13) and for level of return on the safe project satisfying $Z < R_S < B_{CC}$, with Z and B_{CC} defined in (18) and (40), respectively, the bank's investment preference is as follows:

(a) The bank prefers the risky project for $0 \leq c < c_{CC}^{*}$, while for $c_{CC}^{*} \leq c < c^{Sufficient}$ the safe project is preferred, with $c_{CC}^{*} \in (c_{CC}^{Recapitalize}, c^{Sufficient})$, where $c_{CC}^{Recapitalize}$ and $c^{Sufficient}$ are defined in (33) and (10), respectively, and

$$c_{CC}^{*} = 1 - \frac{R_S - pR_H - (1-p-\mu)[R_L(1+\beta) - T]}{\mu - \frac{\beta}{1-c_{\min}}(1-p-\mu)}. \quad (41)$$

(b) The bank prefers the risky project for $c^{Sufficient} \leq c < c_{CC}^{**}$ and the safe project for $c \geq c_{CC}^{**}$, where $c_{CC}^{**} \in (c^{Sufficient}, 1)$ and $c_{CC}^{**} = c^{**}$, with c^{**} defined in (20).

Observe that the introduction of a variable component for recapitalization cost leaves both boundaries of the interval $(c^{Sufficient}, c_{CC}^{**})$ unchanged. Thus, our model is robust to this specification, and a concave cost of recapitalization does not affect the risk-taking incentives of well-capitalized banks when projects exhibiting heavier left tails are available for investment.

References

Acharya, V. V. 2003. "Is the International Convergence of Capital Adequacy Regulation Desirable?" *Journal of Finance* 58 (6): 2745–82.

———. 2009. "A Theory of Systemic Risk and Design of Prudential Bank Regulation." *Journal of Financial Stability* 5 (3): 224–55.

Acharya, V. V., T. Cooley, M. Richardson, and I. Walter. 2009. "Manufacturing Tail Risk: A Perspective on the Financial Crisis of 2007–09." *Foundations and Trends in Finance* 4 (4): 247–325.

Acharya, V. V., L. H. Pedersen, T. Philippon, and M. Richardson. 2010. "Measuring Systemic Risk." Federal Reserve Bank of Cleveland Working Paper No. 1002.

Acharya, V. V., and M. Richardson, eds. 2009. *Restoring Financial Stability: How to Repair a Failed System.* An Independent View from New York University Stern School of Business. Hoboken, NJ: John Wiley & Sons, Inc.

Acharya, V. V., and T. Yorulmazer. 2007. "Too Many to Fail—An Analysis of Time-Inconsistency in Bank Closure Policies." *Journal of Financial Intermediation* 16 (1): 1–31.

Adrian, T., and M. K. Brunnermeier. 2009. "CoVaR." Mimeo, Princeton University.

Angora, A., I. Distinguin, and C. Rugemintwari. 2009. "Excess Capital of European Banks: Does Bank Heterogeneity Matter?" Working Paper.

Asquith, P., and D. W. Mullins, Jr. 1986. "Equity Issues and Offering Dilution." *Journal of Financial Economics* 15 (1–2): 61–89.

Bank for International Settlements. 2010. "Group of Governors and Heads of Supervision Announces Higher Global Minimum Capital Standards." Available at http://bis.org/press/p100912.htm.

Basel Committee on Banking Supervision and Financial Stability Board. 2010. "Assessing the Macroeconomic Impact of the Transition to Stronger Capital and Liquidity Requirements." Interim Report. Bank for International Settlements.

Berger, A., R. DeYoung, M. Flannery, O. Oztekin, and D. Lee. 2008. "How Do Large Banking Organizations Manage Their Capital Ratios?" *Journal of Financial Services Research* 34 (2): 123–50.

Bichsel, R., and J. Blum. 2004. "The Relationship Between Risk and Capital in Swiss Commercial Banks: A Panel Study." *Applied Financial Economics* 14 (8): 591–97.

Blum, J. 1999. "Do Capital Adequacy Requirements Reduce Risk in Banking?" *Journal of Banking and Finance* 23 (5): 755–71.

Boot, A. W. A., and S. I. Greenbaum. 1993. "Bank Regulation, Reputation, and Rents: Theory and Policy Implications." In *Capital Markets and Financial Intermediation*, ed. C. Mayer and X. Vives. Cambridge University Press.

Brunnermeier, M., and L. H. Pedersen. 2009. "Market Liquidity and Funding Liquidity." *Review of Financial Studies* 22 (6): 2201–38.

Caminal, R., and C. Matutes. 2002. "Market Power and Banking Failures." *International Journal of Industrial Organization* 20 (9): 1341–61.

De Jonghe, O. 2010. "Back to the Basics in Banking? A Micro-Analysis of Banking System Stability." *Journal of Financial Intermediation* 19 (3): 387–417.

Dewatripont, M., and J. Tirole. 1994. *The Prudential Regulation of Banks*. Cambridge, MA: MIT Press.

Flannery, M. 1989. "Capital Regulation and Insured Banks' Choice of Individual Loan Default Risks." *Journal of Monetary Economics* 24 (2): 235–58.

Furlong, F., and S. Kwan. 2005. "Market-to-Book, Charter Value, and Bank Risk-Taking — A Recent Perspective." Working Paper, Federal Reserve Bank of San Francisco.

Gorton, G. 2010. *Slapped by the Invisible Hand: The Panic of 2007*. New York: Oxford University Press.

Hellmann, T. F., K. C. Murdock, and J. E. Stiglitz. 2000. "Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?" *American Economic Review* 90 (1): 147–65.

Holmstrom, B., and J. Tirole. 1997. "Financial Intermediation, Loanable Funds and the Real Sector." *Quarterly Journal of Economics* 112 (3): 663–91.

Huang, R., and L. Ratnovski. 2009. "Why Are Canadian Banks More Resilient?" IMF Working Paper No. 09/152.

———. 2011. "The Dark Side of Bank Wholesale Funding." *Journal of Financial Intermediation* 20 (2): 248–63.

International Monetary Fund. 2010. "United States: Publication of Financial Sector Assessment Program Documentation—Technical Note on Stress Testing." IMF Country Report No.10/244 (July).

Jensen, M. C. 1986. "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers." *American Economic Review* 76 (2): 323–29.

Jensen, M. C., and W. H. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3 (4): 305–60.

Kahane, Y. 1977. "Capital Adequacy and the Regulation of Financial Intermediaries." *Journal of Banking and Finance* 1 (2): 207–18.

Keeley, M. C. 1990. "Deposit Insurance, Risk, and Market Power in Banking." *American Economic Review.* 80 (5): 1183–1200.

Kim, D., and A. M. Santomero. 1988. "Risk in Banking and Capital Regulation." *Journal of Finance* 43 (5): 1219–33.

Knaup, M., and W. Wagner. 2010. "Measuring the Tail Risks of Banks." Working Paper.

Koehn, M., and A. M. Santomero. 1980. "Regulation of Bank Capital and Portfolio Risk." *Journal of Finance* 35 (5): 1235–44.

Matutes, C., and X. Vives. 2000. "Imperfect Competition, Risk Taking, and Regulation in Banking." *European Economic Review* 44 (1): 1–34.

Myers, S., and N. Majluf. 1984. "Corporate Financing and Investment Decisions When Firms Have Information that Investors Do Not Have." *Journal of Financial Economics* 13 (2): 187–221.

Perotti, E., and J. Suarez. 2009. "Liquidity Risk Charges as a Macroprudential Tool." CEPR Policy Insight No. 40.

———. 2010. "A Pigovian Approach to Liquidity Regulation." CEPR Discussion Paper No. 8271.

Repullo, R. 2004. "Capital Requirements, Market Power, and Risk-Taking in Banking." *Journal of Financial Intermediation* 13 (2): 156–82.

Rochet, J.-C. 2004. "Macroeconomic Shocks and Banking Supervision." *Journal of Financial Stability* 1 (1): 93–110.

Shin, H. S. 2009. "Reflections on Northern Rock: The Bank Run that Heralded the Global Financial Crisis." *Journal of Economic Perspectives* 23 (1): 101–19.

Discussion of "Capital Regulation and Tail Risk"

Andrew Winton
Carlson School of Management, University of Minnesota

During the recent financial crisis, the notion of "tail risk"—exposure to very unlikely yet massive losses—rapidly became the foremost concern of regulators, banks, and other market participants alike. Perotti, Ratnovski, and Vlahu (this issue) analyze how the presence of such risks affects the relationship between bank capital and bank risk taking. With policymakers looking to revamp capital regulations so as to prevent a similar crisis from occurring, there is no question that this is a very timely paper indeed.

Perotti, Ratnovski, and Vlahu argue that tail risk differs from the "normal" risks modeled in the banking literature because it can wipe out any amount of bank capital. As a result, the presence of tail risk weakens the effectiveness of capital in reducing risk-shifting incentives. Moreover, when capital is costly to raise, the combination of tail risk and less-catastrophic ("non-tail") risk can complicate the relationship between capital and risk shifting, making it non-monotonic: although an increase in bank capital from low levels tends to deter risk shifting, at even higher levels the bank may start taking risks again, only to revert to a safer strategy if capital is sufficiently high. This possible reversal arises from the fact that regulatory capital requirements apply not only ex ante but also to the bank's condition after risks have materialized. The reversal becomes stronger as the probability of tail risks increases. The upshot is that capital regulation is even less effective in the presence of tail risk.

This paper makes a number of innovations, including not only the distinction between tail and non-tail risk but also the impact of risk realizations on ex post capital adequacy and how, in the presence of recapitalization costs, this in turn feeds back into initial risk choices. Nevertheless, I think there are several aspects in which it could be improved. The result that tail risk makes capital less effective is not unique to this model. Also, the way in which tail risk affects the bank's risk-shifting opportunities is somewhat at odds with the factors that helped lead to the recent financial crisis. The robustness and applicability of capital's non-monotonic effects on risk shifting could be explored in more depth. Finally, some examination of the impact of tail risk on social welfare and optimal capital regulation would greatly increase the paper's contribution. Before discussing these points at more length, I will summarize the authors' model and results.

1. Summary

In the model, there are three dates. At date 0 a risk-neutral bank begins with exogenous capital C and insured deposits D, and chooses between safe and risky assets. At date 1, date 2 asset returns become known to all. The bank must meet a minimum capital ratio $c_{\min} > 0$ or be closed, with bank shareholders receiving nothing in the event of closure; if it chooses to, the bank can raise additional capital at a fixed cost T. Finally, at date 2, the asset returns are realized, depositors are paid, and bank shareholders receive any surplus.

Safe assets yield a constant gross return $R_S > 1$ at date 1. Risky assets have a yield that depends on the state of the economy at date 1: they return $R_H > R_S$ with probability p (the high-return state); they return R_L, where $1 > R_L > 0$, with probability $1 - p - \mu$ (the low-return state); and they return 0 with probability μ (the tail-return state). The expected return on the safe asset is higher than that on the risky asset, so the safe asset is the efficient choice.

Following well-known arguments, sufficiently low bank capital levels will give the bank incentive to prefer the risky asset even if there is no cost to raising capital (proposition 1). If the safe return R_S is too high, this situation only occurs if capital is very low—low enough that the bank would default even in the low-return state. Otherwise, the bank may prefer the risky asset even if the bank would only default in the tail-return state. The rest of the paper concentrates on the latter case.

Comparative statics show that a mean-preserving increase in tail risk causes the initial minimum capital level required to prevent risk shifting, $c^{T=0}$, to increase at a rate that is inversely proportional to the square of the probability μ that the tail-return state occurs. Perotti, Ratnovski, and Vlahu interpret this as showing the diminished effectiveness of capital at preventing risk shifting when the bank has access to assets with tail risk.

The authors then turn to the case where raising additional capital at date 1 is costly. Now, two effects come into play in determining how initial capital levels affect the bank's risk-taking incentives. The first is the deterrent effect already noted: with more initial capital, the bank has more to lose in the event of realized tail risk, so risk shifting is less attractive. The second is new: more capital means that the bank is less likely to have to raise additional capital to meet the required capital level $c_{\min}$ in the low-return state. (Recall that, by assumption, the bank does not default in this state, so capital exceeds zero, but it might still fall short of $c_{\min}$.)

Proposition 2 establishes that if the return on the safe asset is neither too high nor too low, the bank's asset choice depends on which of four regions its initial capital level falls into.[1] For low capital, the bank takes the risky asset; for somewhat higher capital, it chooses the safe asset so as to avoid having to recapitalize or abandon its equity position in the low-return state. Nevertheless, for even higher capital levels, the bank will choose the risky asset because it does not have to recapitalize in the low-return state, making the net effect of risk shifting (in the tail-return state) positive. Finally, for sufficiently high initial capital, even risk shifting through the tail-return state is unattractive, and the bank chooses the safe asset. Mean-preserving increases in tail risk expand the third region, where the bank takes on tail risk because it has enough capital to avoid recapitalization in the low-return state.

2. Remarks

This paper makes a nice contribution by showing that, in some cases, costs of recapitalization can make higher bank capital levels have a counterintuitive effect, increasing risk-shifting incentives. As Perotti, Ratnovski, and Vlahu note in their introduction, others

[1] If the return on the safe asset is either sufficiently low or sufficiently high, there are only two relevant capital regions, with the risky asset being chosen if initial capital is low and the safe asset being chosen if initial capital is high.

have shown that higher bank capital requirements can have counterintuitive effects on risk-taking behavior. What distinguishes the current paper is the emphasis on higher *initial capital levels* (as opposed to capital requirements) and their interaction with future return realizations and recapitalization in the presence of tail risk. Nevertheless, the contribution is somewhat less than the authors claim, and more could be done to establish and extend the results.

First, Perotti, Ratnovski, and Vlahu's interpretation of proposition 1 and the comparative statics results that follow is somewhat misleading. That lower returns on the safe asset increase the initial capital needed to deter risk shifting is not new, nor is it new that increases in the probability of (very) bad returns increase the amount of initial capital needed to deter risk shifting. Also, the finding that mean-preserving increases in the probability μ of tail returns cause the risk-deterring capital level $c^{T=0}$ to rise at a rate proportional to μ^{-2} actually suggests that, as tail risk becomes more likely, less additional capital is needed to restore incentives. This is not surprising, since higher initial bank capital means that less of any tail loss is shifted to depositors, so that a further increase in the probability of this happening has a smaller impact on the attractiveness of risk shifting. Finally, all of these results can be obtained in a model with only two return realizations, undercutting the notion that tail risk differs from the impact of low returns in general.[2]

Another issue has to do with the nature of the risk-shifting choice facing the bank. Taking the model literally, a completely safe asset has to have a positive net return that also exceeds the expected return on an asset with exposure both to low returns and to tail risk. One difficulty here is that there are relatively few completely safe assets with positive net present value (NPV); most of the assets banks hold that are likely to have positive NPV also have some risk exposure. Indeed, one partial cause for the recent crisis was institutions' desire for safe AAA assets with above-normal returns—a desire that was met with securities that proved to have considerable tail risk. This suggests that a more realistic and topical approach to

[2]In particular, if case (b) of proposition 1 holds, so that the bank fails in the low-return state as well as the tail-return state, then increases in the combined probability $1 - p$ of either state occurring cause the risk-deterring level of capital to increase at a rate proportional to the inverse square of this probability.

tail risk would be to give the bank a choice between "risky" assets that return $R_H > 1$ in the high-return state and $R_L < 1$ in both the low-return and tail-return states, and "safe" assets that return $R_S > 1$ in the high- and low-return states and 0 in the tail-return state. Examining how capital levels affect this choice might yield interesting insights into the recent crisis and how to prevent a recurrence.

Turning to the results with costs of raising capital, my first comment concerns robustness. Perotti, Ratnovski, and Vlahu show that the results are robust to the bank having a franchise value that is lost on default and to having both fixed and variable costs of raising capital. A natural question is the extent to which the results depend on having three discrete states, with capital shortfalls in the low-return state traded off against potential risk shifting in the tail-return state. Intuitively, if returns are continuously distributed, there will still be regions where capital is wiped out, where recapitalization is an issue, and where recapitalization is not necessary. It seems reasonable that a riskier asset that shifts the return density in the right way will lead to similar results on the effects of capital increases, but it would be nice to work out just what sort of density shift is needed and how likely this is to occur in practice. My previous comments about the nature of tail-risk assets seem applicable here as well.

A related issue has to do with the timing of recapitalization and resolution of uncertainty. In the model, the recapitalization decision occurs after future returns are known with certainty; yet, in reality, policymakers and researchers have been concerned with residual uncertainty at the time of the recapitalization decision and the debt overhang problem this creates. Some discussion of how residual uncertainty and debt overhang are likely to affect the model's results would be useful.

Another facet of applicability has to do with the parameter restrictions that are required for the results of interest: those that guarantee risk shifting only takes place through tail risk, and those that guarantee an interesting trade-off between risk-shifting concerns in the tail-return state and recapitalization concerns in the low-return state. In order to assess the importance of these results, it would be helpful to know what real-world situations correspond to these parameter restrictions and how likely they are to occur in practice.

Finally, the analysis takes initial capital and capital requirements as costless and exogenous and ignores questions of how tail risk affects the broader economy. Since even initial levels of capital may involve private and social costs (Diamond and Rajan 2000, Gorton and Winton 2000, and Hellmann, Murdock, and Stiglitz 2000), and regulatory responses during the crisis focused on possible externalities from tail-risk-induced bank failure, incorporating some notion of these costs of capital and bank failure and their effect on social welfare would greatly increase the paper's impact.

3. Conclusion

Perotti, Ratnovski, and Vlahu demonstrate that recapitalization costs can have counterintuitive effects on the relationship between bank capital and risk shifting, particularly in the presence of tail risk. Although I have suggested some ways in which the analysis can be extended and deepened, the paper is intriguing and the topic is well worth further study.

References

Diamond, D., and R. Rajan. 2000. "A Theory of Bank Capital." *Journal of Finance* 55 (6): 2431–65.

Gorton, G., and A. Winton. 2000. "Liquidity Provision, Bank Capital, and the Macroeconomy." Working Paper.

Hellmann, T., K. Murdock, and J. Stiglitz. 2000. "Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?" *American Economic Review* 90 (1): 147–65.

The Capital Conundrum*

Richard J. Herring
The Wharton School, University of Pennsylvania

After a review of the theory of regulation of bank capital, this paper notes that the pervasive influence of the safety net provides both a rationale for regulating bank equity capital and an obstacle to inferring what the optimal capital-to-asset ratio would be for a bank in the absence of the safety net (or expectations of an ex post bailout). This paper supports the view that the cost of bank equity capital is less than is frequently assumed, but notes that many of the frictions that lead to optimal equity-to-asset ratios for other firms are likely to apply to banks. Moreover, the analysis of bank capital structures is further complicated by the fact that a significant proportion of bank liabilities—deposits—are an important product offered by banks as well as a means of increasing leverage. After a brief overview of the potential advantages of a requirement for contingent convertible capital (CoCo) instruments in addition to higher equity-to-asset ratios, the paper argues that, given the uncertainty about the optimum equity capital requirement, a substantial CoCo requirement provides additional advantages, which include stronger incentives for banks to recapitalize before they encounter serious difficulties, enhanced incentives for banks to adopt the best possible risk-management measures, and (so long as the regrettable asymmetry between interest and dividends remains) reduced incentives for banks to move activities to the shadow banking system. A substantial CoCo requirement protects society from loss as effectively as an equivalent amount of additional equity capital, but CoCos enable a bank to recapitalize automatically if it falls short of the equity capital requirement. This recapitalization will occur instantaneously and at lower

*The author is grateful to participants in the Third Financial Stability Conference of the International Journal of Central Banking in London and to Til Schuermann for comments and to Tina Horowitz and Christopher Trollen for invaluable assistance in producing this paper.

cost than a new issue of equity under conditions of stress. Instantaneous recapitalization will give the bank an opportunity to restructure or find a private solution and will provide the regulatory authorities with sufficient warning to prepare a rapid resolution if necessary.

JEL Codes: G21, G28, and G32.

Rigorous academic thinking about capital structure began with the seminal article by Miller and Modigliani (M&M) (1958). A quartet of eminent financial economists (Admati et al. 2011) have argued that, with regard to bank capital regulation, M&M was not only the first word, but should also have been the last. They have expanded, in considerable detail, some of the arguments Merton Miller (1995) presented at a Wharton conference in 1994.

At root the argument is very simple and intuitively appealing. In the frictionless world assumed by M&M, a firm's choice of capital structure cannot affect the value of its assets. Thus, in the context of bank regulation, where default of a large, complex financial institution is assumed to have significant spillover costs on the rest of society, banks should be required to issue a very high proportion of equity capital without loss of value. This high proportion of equity may seem absurd to some, but the capital structure of mutual funds is 100 percent equity and this sector of the financial services industry holds assets that exceed 80 percent of the assets of U.S. commercial banks (Federal Reserve Board 2011). It differs markedly, however, from the capital structure banks have chosen (subject to regulatory constraints).

Til Schuermann (2011) has made the tongue-in-cheek observation that the wide dispersion of equity-to-asset ratios across firms—varying from single digits to 100 percent—suggests that M&M may not be wrong. The choice of capital structure among firms may be essentially a random decision without impact on the value of the firm. When firms are grouped into industries, however, a different pattern emerges. (See table 1.) Although some dispersion exists among firms within a particular industry, it is less than the dispersion found across industries. Moreover, even though leverage varies markedly across industries, industries earn remarkably similar

Table 1. Profitability and Leverage across Industries (Medians across Years and Institutions)

	Return on Assets[a]				Return on Equity[b]				Leverage[c]			
	95–09	95–00	01–07	08–09	95–09	95–00	01–07	08–09	95–09	95–00	01–07	08–09
Banks	0.6	0.7	0.7	0.2	12.2	13.3	12.8	3.2	18.3	17.8	19.1	17.4
Non-Bank Financials	0.9	1.0	1.0	0.5	11.2	12.3	11.4	5.4	12.1	12.5	12.1	10.8
Non-Financials	3.2	3.0	3.4	2.8	11.7	10.9	12.8	9.8	3.0	3.0	3.0	2.9
Energy	5.9	3.9	8.1	5.2	14.2	10.8	18.6	10.1	2.4	2.5	2.3	2.2
Materials	4.3	4.3	4.7	3.2	10.6	8.8	13.1	8.5	2.5	2.4	2.5	2.7
Industrials	2.1	1.4	2.4	2.3	10.4	8.3	11.5	11.0	5.4	6.1	5.4	4.8
Consumer Discretionary	2.2	2.1	2.6	1.1	9.1	8.9	10.4	4.2	3.4	4.0	3.1	3.1
Consumer Staples	5.4	5.2	5.7	5.1	13.0	12.4	13.8	11.7	2.5	2.4	2.5	3.0
Health Care	8.1	8.0	8.3	6.5	18.2	18.8	18.5	15.3	2.3	2.3	2.3	2.3
Information Technology	5.1	5.1	5.0	5.6	12.8	15.1	12.8	10.3	2.2	2.2	2.1	2.0
Telecom Services	3.2	3.6	2.8	2.9	8.5	10.8	8.4	6.4	2.6	2.7	2.6	2.7
Utilities	2.7	2.5	2.7	2.7	10.8	9.3	11.6	11.9	4.1	3.7	4.4	4.0

[a]Net income over total assets, in percent.
[b]Net income over total shareholder funds, in percent.
[c]Total assets over total shareholder funds.
Source: Bloomberg.

returns on equity. It is also notable that among all industries, banks have the highest leverage.

A considerable amount of work in corporate finance since M&M has focused on the consequences of introducing a variety of frictions into their frictionless world to determine which frictions are most likely to explain systematic differences in optimal capital structures across industries. One obvious friction derives from an asymmetry found in most national tax codes. Interest costs may be deducted from taxable income, but dividends cannot. Thus, other things equal, a firm can increase value to its shareholders by increasing its leverage. Of course, this is a private benefit, not a social benefit, and should not enter into a regulator's computation of the optimal equity-to-asset ratio, except that it is important to recognize the distortions this creates and the unintended consequences that may arise when treating banks differently from other firms.

Other frictions derive from a variety of asymmetric information and agency costs. One set of frictions can be categorized as the costs of financial distress. If bankruptcy is costly—and recent evidence from the ongoing bankruptcy proceedings for Lehman Brothers suggests that it is—then any entity conducting transactions with the firm will try to avoid incurring such costs. Since a firm's probability of default rises as its leverage increases, beyond some point the probability of default will become sufficiently high that a firm's costs will start to increase. This is not only because a firm's probability of bankruptcy has risen to worrisome levels, but also because the firm may be tempted to exploit asymmetric information by engaging in asset substitution and taking riskier bets. These costs of financial distress include not only the cost of borrowing (and often the lack of availability of borrowing) but also the costs of various other inputs. Suppliers are likely to demand cash on delivery—or even cash in advance (an increasingly important concern as outsourcing has expanded)—and a firm must also worry about employee costs. Some of the most talented employees will leave for more secure positions, and they must be replaced by new employees who are likely to be less efficient for a considerable period of time. Moreover, it may be necessary to offer "golden handcuffs" to key personnel deemed vital to the functioning of the firm.

These costs of financial distress lead a firm to prefer a more moderate degree of leverage than if only the tax advantages of debt were taken into account. Because these costs of financial distress are likely to vary across industries, it should not be surprising to observe different industries clustering around different equity-to-asset ratios. Of course, this analysis can be made much more sophisticated, but the important point is that these arguments apply to all firms. So far, none of these frictions pertains uniquely to banks.

In general, the equity-to-asset ratio of non-financial industries is not a matter of public concern. In most such industries, we assume that creditors and shareholders internalize most of the costs of bankruptcy and so market forces will lead firms to adopt optimum capital ratios that take into account the particular frictions they face.

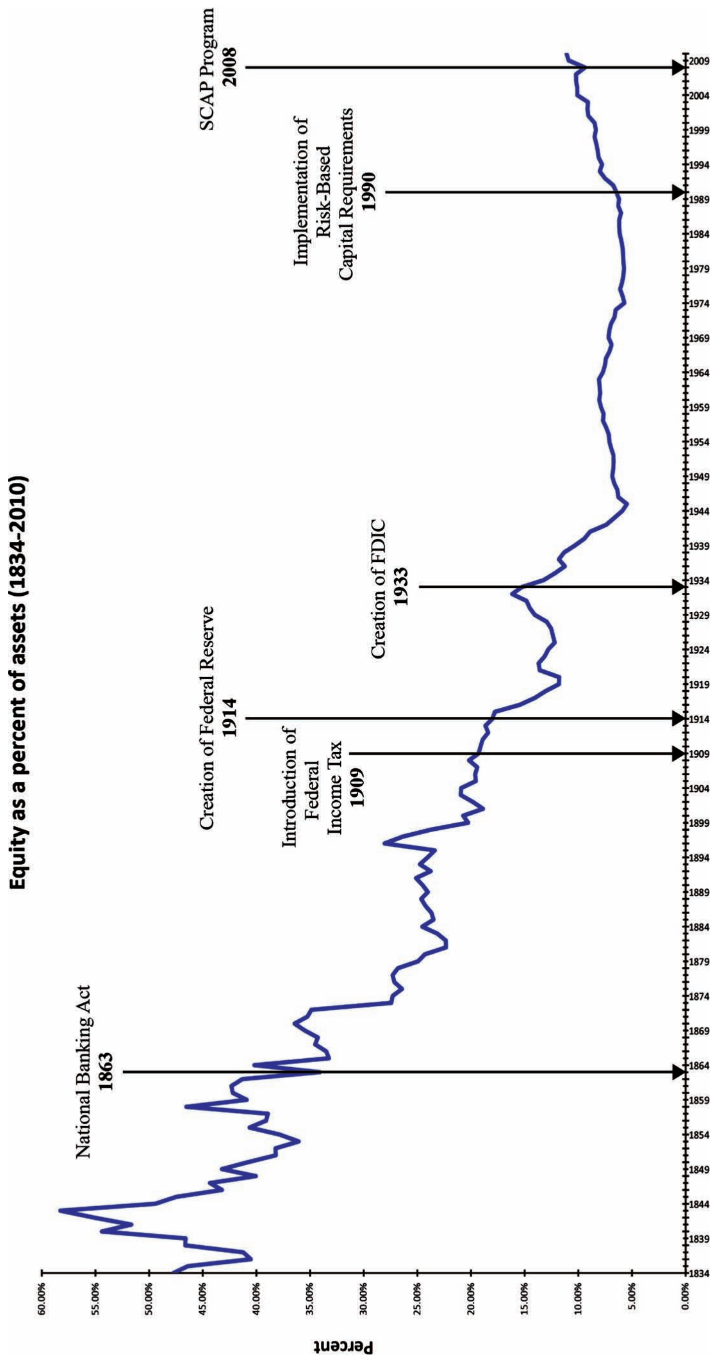
This argument does not apply to the banking industry, however, because virtually every country has erected a safety net to protect the financial system and, more broadly, the economy, from damaging spillovers from a banking crisis—costs that are not internalized by the creditors and shareholders of banks when they take decisions regarding the bank's optimal capital ratio. Although the details of the safety net differ from country to country, they generally involve implicit guarantees for the asset side of a bank's balance sheet (primarily through lender-of-last-resort facilities) and explicit, as well as implicit, guarantees for the liability side of the balance sheet.

Deposit insurance constitutes the principal explicit guarantee (which, as we have seen, may be expanded greatly in a crisis). Implicit guarantees can arise from the practice of relying on "purchase and assumption" transactions to resolve a failing bank or simply a record of protecting uninsured creditors in the event of failure. The safety net also includes financial supervision, which is intended to assure the public that banks are following the rules designed to ensure their safety and soundness. If the public has confidence in the safety net, the expected costs of financial distress fall so that banks will prefer substantially greater amounts of leverage than they would choose if the safety net were not available. This, indeed, is the principal rationale for imposing capital requirements on banks.

By looking back to the early nineteenth century, we can make some inferences about the degree of leverage banks would prefer in the absence of various elements of the safety net.[1] Figure 1 shows

[1] This extends the discussion found in Berger, Herring, and Szegö (1995).

Figure 1. The Evolution of Leverage in U.S. Banks (1834–2010)



the evolution of equity-to-asset ratios for U.S. banks from 1834 to the present.[2] The pattern is striking. Before 1863, no federal banking regulation existed. Banks did not enjoy access to any of the elements of a safety net and they chose very high equity-to-asset ratios because the costs of financial distress were very high. These could not be offset against the tax benefits of debt because corporate income taxes did not exist. These ratios, which were as high as 55 percent, declined markedly with the enactment of the National Banking Act of 1863, which created the Office of the Comptroller of the Currency (OCC). The OCC chartered and supervised national banks. To the extent that depositors and creditors had confidence in delegating monitoring to the OCC, the introduction of this element of the safety net may have reduced their concerns about asymmetric information. Equity-to-asset ratios fell quite dramatically to a range between 30 percent and 35 percent.

The remaining decline in equity-to-asset ratios over the nineteenth century can be largely attributed to technological and institutional improvements that reduced the vulnerability of many banks to a run. The introduction of the telegraph, national railway connections, regional exchanges, and bank clearinghouses fostered the development of regional and embryonic national financial markets, increasing the access of banks to sources of liquidity beyond their own markets.

Just after the turn of the century, two policy measures may have induced banks to assume higher leverage, but they happened so quickly in succession that it is difficult to disentangle their separate contributions. The corporate income tax was introduced in 1909, but initial tax rates were so low that they are unlikely to have led to significant distortions. Just five years later, the Federal Reserve System was created with the power to engage in discount window lending. Shortly after World War I began, corporate tax rates were raised to levels that may well have biased the capital structure choice in favor of greater leverage. Probably for both reasons, leverage began to rise until the start of the Great Depression, when heightened creditor and depositor concerns about asymmetric information led banks

[2] Since accounting conventions have shifted over time, these ratios must be regarded as a rough indication of trends.

to deleverage. This was only a brief interruption of a century-long trend of increasing leverage.

The introduction of explicit deposit insurance, with the creation of the FDIC in 1933, led to an increase in leverage—with equity-to-asset ratios falling to the 5 percent to 10 percent range, where they remained until the introduction of the Basel I requirements in 1990 and the Federal Deposit Insurance Improvement Act (FDICIA) leverage ratio in 1991. Among other things, FDICIA introduced a structured early-intervention system tied to increases in leverage below the required equity-to-asset ratio. FDICIA was probably more important in leading U.S. banks to decrease their leverage than the Basel Accord because the leverage standard frequently required more equity capital than the Basel risk-weighted capital requirements. From that point, the evolution of leverage among U.S. banks differed from that in most other countries. Banks subject only to the Basel risk-weighted ratios tended to increase their leverage as they found ways to arbitrage the risk weight, while banks in the United States were constrained by the leverage requirement.

The final reduction in leverage for U.S. banks was a direct consequence of measures taken to restore confidence in the wake of the crisis. In 2009, the U.S. regulatory authorities introduced the Supervisory Capital Assessment Program (SCAP) that required banks to project their capital positions under stress conditions. Nineteen of the largest institutions were required to conduct simulations to determine whether their capital buffers were sufficient to withstand the specified degree of stress. Ten of the nineteen were deemed to have failed the test even though they met the legally mandated capital requirement. That this approach strengthened rather than weakened confidence can be explained by the fact that banks that failed were required to accept a capital infusion from the government until they could reduce their assets or raise sufficient capital on their own.

It is evident that the Basel II approach to regulating bank capital failed comprehensively. The numerator in the regulatory ratio did not reflect an institution's ability to absorb loss without going through resolution or bankruptcy, and the denominator did not capture the most important risks to which banks were exposed. Moreover, the minimum was set much too low, which is not surprising since no rationale for the original 4 percent tier 1 ratio and the 8 percent combined tier 1 and tier 2 ratio has ever been offered by the

Basel Committee.[3] This left bank supervisors all over the world in the awkward situation of trying to explain why it was necessary to provide public funds to a bank that met and even exceeded the minimum regulatory capital requirement. Indeed, banks that required intervention often reported higher regulatory capital ratios in the preceding period than other banks that did not require assistance.

Even though systemic banking crises can be enormously expensive—for example, by June 2009, the United States and Europe had committed roughly 25 percent of world GDP to guaranteeing their banking systems (Alessandri and Haldane 2009)—many legislators, regulators (and, of course, bankers) appear to believe that banks add sufficient value in terms of payment services, intermediation, and maturity transformation to warrant maintenance of a high degree of leverage regardless of the potential costs. This presumption lacks rigorous empirical support. In fact, academics and researchers in some regulatory institutions are undertaking increasingly sophisticated studies to attempt to quantify the trade-offs that may be involved in requiring banks to fund themselves with substantially more equity capital.

In the meanwhile, regulators have responded to the crisis by proposing that internationally active banks be subject to equity capital requirements as much as four to five times current (very low)[4] international minimum risk-adjusted ratios (Basel Committee on Banking Supervision 2011). Some regulators (Tarullo 2011), several academics, and even the *Wall Street Journal* (2011) have argued for still higher equity capital requirements.

[3]The most plausible rationale is the cynical observation that the 4 percent/8 percent standard was set so as not to inconvenience any major bank.

[4]Tier 1 capital was originally intended to be mainly equity, retained earnings, and instruments such as non-cumulative, perpetual preferred stock that could serve as a source of strength to maintain the bank as a going concern. However, over time, pressures from banks and creative investment bankers led to the acceptance of a number of hybrid instruments that appeared to be sufficiently like equity to placate the regulators and sufficiently like debt to convince the tax authorities that payments on such instruments could be counted as interest payments and deducted in computing taxable income. Thus, over time, what was originally a 4 percent equity requirement fell to a 2 percent equity requirement. Most of these hybrid instruments proved utterly useless to sustain the bank as a going concern because they could bear loss only in a resolution or bankruptcy process. Regulators intend to ban these instruments in the new definition of tier 1 capital that they hope to implement by 2018.

Banks have countered that higher equity capital requirements will require that they charge correspondingly higher spreads to customers, which would cause a decline in economic growth and shift a considerable amount of traditional bank business to the largely unregulated shadow banking sector (Institute of International Finance 2010). Although the relationship between spreads and the rate of return on equity seems clear in the mechanical formula sometimes used to relate spreads to the equity-to-asset ratio,[5] this result is based on the assumption that *required* returns on equity and funding costs will not adjust to changes in a bank's leverage.

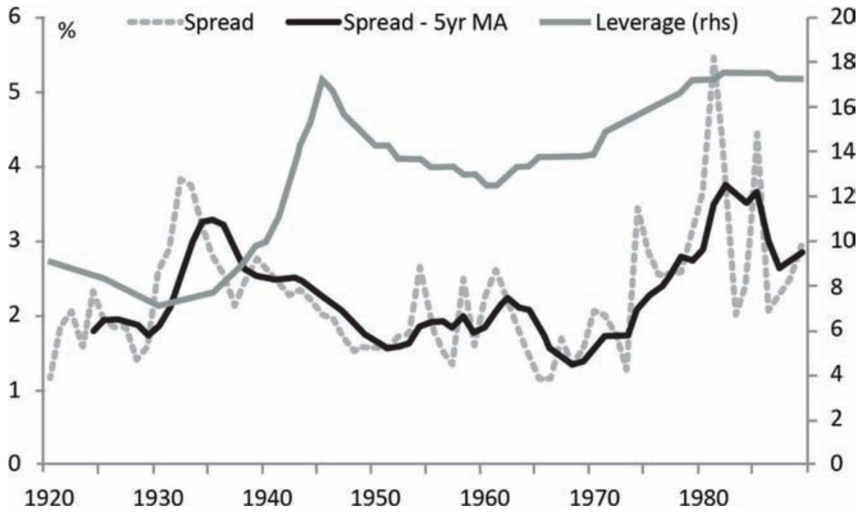
Of course, the logic of the M&M argument suggests that the market would be willing to pay more for a stream of less-leveraged earnings (because it is less risky and therefore the required return on equity will decline) and creditors will charge lower credit-risk premia (because the bank is less likely to default). The distortions created by the safety net, however, imply that banks will not experience the full benefit of a reduction in leverage because many creditors are already relying on external guarantees provided by government rather than the bank's own creditworthiness. But the implicit public subsidy is reduced. In addition, Miles, Yang, and Marcheggiano (2011) present evidence from 1880 to this decade showing that spreads have borne no clear relationship to bank leverage in the United States. (See figure 2.) Moreover, they present additional evidence from the United Kingdom showing that leverage has not had a clear impact on the growth of real GDP.

We face a genuine conundrum, however, in determining the optimal equity-to-asset ratio for banks. We can't rely on market data because they are hopelessly distorted by the numerous features of the safety net (net of regulatory compliance costs). The calculation is further complicated by the fact that deposit liabilities, which provide a store of value and transactions services, are a key part of the

---

[5]This relationship can be derived from the plausible assumption that banks must set spreads at least large enough to yield the required return on equity that shareholders demand to avoid destroying shareholder value. This proposition can be reduced to the simple formula that spreads must be set at least as high as $S = E/A\ (r - i)$, where S is the spread above the cost of funds, E is equity, A is total assets, r is the required return on equity, and i is the bank's cost of funds. If one assumes that $(r - i)$ will not change, higher equity-to-asset ratios require correspondingly higher spreads.

Figure 2. Leverage Has Not Had a Clear Impact on Spreads



Source: Homer and Sylla (1991)
Note: This figure is as shown in Miles, Yang, and Marcheggiano (2011, p. 8).

value that banks add to society. Since some liabilities are really a product supplied by the bank rather than simply a means of funding the bank, we know that a 100 percent equity-to-asset ratio cannot be the correct answer. But given the enormous costs incurred by governments on both sides of the Atlantic, it is clear that equity-to-asset ratios have been set much too low to provide a sufficient buffer against the risks banks have taken. Thus, as most regulators recognize, there is a strong case for requiring much higher equity-to-asset ratios.

The leverage ratio contained in the Basel III proposal, however, is quite timid—a mere 3 percent—which would have been inadequate to enable several major banks to continue as going concerns without government assistance. It seems clear that the equity-to-asset ratio should be raised substantially above the level contemplated in the current proposal. But how much higher? We lack a clear analytic answer to this important question, but one ballpark estimate might be equal to the amount of equity and long-term debt banks currently issue[6]—unless there is a case for supplementing an equity-to-asset ratio with a requirement for an additional contingent convertible debt ratio to provide a means of efficiently recapitalizing a bank if a shock reduces its equity-to-asset ratio below its required level.

Calomiris and Herring (2011)[7] have argued that there is a strong case to be made for requiring much higher equity-to-asset ratios and for restating these regulatory requirements in terms of proxies for market values rather than relying exclusively on accounting measures which can be easily manipulated and will always lag the actual deterioration in a bank's net worth (Herring, forthcoming). Moreover, we argue that this greatly strengthened equity-to-asset ratio should be supplemented with a required ratio of properly structured convertible contingent bonds (CoCos).

What can CoCos accomplish that cannot be accomplished by an equity-to-asset ratio alone? Some proposals for CoCos emphasize their role as a buffer against loss in the event a bank must be resolved. Others emphasize that CoCos can be a useful signaling device and a way to supplement supervisory discipline with market discipline. While we recognize these potential benefits, we think that CoCos can be designed to accomplish an even more important function. Our proposal is designed to provide strong incentives for managers to maintain high equity-to-asset ratios and strong risk-management controls. If managers fail to do so, shareholders will be substantially diluted and the managers will almost certainly be replaced by new management because both old and new shareholders are likely to be displeased with their performance. In addition, CoCos will provide a way for a bank that falters to reestablish its equity-to-asset ratio without having to go to the market at an unfavorable time (when the "lemons problem" is likely to be most severe) and to enable the bank to maintain its required equity-to-asset ratio without abruptly reducing its assets, which can destabilize markets and reduce the availability of credit when it may be most needed. In short, we believe that a requirement for appropriately structured CoCos (alongside a substantially higher required equity-to-asset

[6]This is because long-term debt is generally issued simply to fund the bank rather than to fulfill customer demands.

[7]This builds on the pioneering work by Flannery (2005) that introduced the notion of contingent convertible bonds in the banking literature.

ratio) will be a more effective prudential tool than a higher equity requirement that is equal to the sum of our increased equity-to-asset ratio and our CoCo-to-asset requirement.

What differentiates our proposal from several other variants is the amount of CoCos that we would require that banks issue, the trigger for conversion of CoCos into equity, the quantity of CoCos converted when the trigger is set off, and the price at which debt is converted into equity. These features are designed to make the conversion of CoCos so costly to shareholders and managers that they will take every possible precaution to avoid triggering the conversion. This is also important in expanding the market for CoCos. If the chance that CoCos will be converted is sufficiently low, they will be priced like subordinated debt and will appeal to a much broader range of institutional investors.[8]

We advocate setting regulatory capital requirements with regard to the quasi market value of assets (QMVA), which we define as the market value of equity plus the face value of debt. This avoids the difficult problem of making a timely evaluation of a bank's assets (which are often illiquid and opaque) by taking advantage of the information contained in equity prices.[9] Moreover, we believe that capital requirements should be set under the assumption that the bank will continue as a going concern and so liabilities should be valued at face value, which is easily observed.[10]

[8]Higher equity-to-asset ratios will, of course, reduce the price of all debt in principle, but the point is that a sufficiently low probability of conversion should make the instrument appealing to many of the same institutions that hold long-term claims on banks.

[9]In this era of highly volatile share prices, some observers doubt that there is significant information content in market values, and we would certainly not want to rely on a bank's share price on any given day, but we show that a reasonable equity requirement based on a 90-day moving average of equity prices (to reduce the impact of day-to-day volatility) would have done an excellent job of distinguishing banks that would require significant government assistance or resolution from those that did not. Moreover, regulators would have been able to see the deterioration several weeks, and in some cases months, before intervention was necessary, thus reducing pressure to design hasty rescue packages over sleepless weekends (see Calomiris and Herring 2011).

[10]Ideally, regulators might want to adjust liabilities for changes in interest rate risk, but *not* for declines in the value of a bank's liabilities due to market perceptions of increased credit risk.

We believe the amount of CoCos issued should be a significant proportion of equity so that management and shareholders will need to focus on the possibility of a significant dilution of the value of their shares if they fail to maintain adequate capital ratios or manage risks carefully. For the same reason, we would argue that the full amount of CoCos should be converted and the conversion price should be very favorable to holders of CoCos—perhaps as high as 1.05 times the face value of their claims.

The trigger for a conversion of CoCos should be a *sustained* decline in the moving average market cap of the bank to the QMVA. We term this ratio the "quasi market value of equity ratio" (QMVER). The trigger should be set at a sufficiently high level so that the market value of equity is not contaminated by either the option value of equity or the expectation of a bailout and so that there is sufficient time for a bank to take corrective action.

In effect we've tried to design a security that would reverse the perverse incentives of a debt overhang in which shareholders are reluctant to issue new equity even though their bank is undercapitalized because most of the increase in value would go to creditors. Under our plan, shareholders would have heightened incentives to issue new equity, take corrective action, or sell the firm before they hit the trigger point because they face the prospect of substantial dilution and, in the case of managers, job loss, if they do not.

Of course, these incentives will not prevent every bank from triggering the conversion of the CoCos. But even when a bank hits the trigger point, society will be better protected from loss than if the CoCos had not been in place. The bank will automatically be recapitalized without incurring the very heavy transaction costs of issuing new equity under pressure when concerns about asymmetric information are most intense. This additional equity may give the (probably new) management team time to restructure the bank or merge the bank with another institution.

Inevitably, some banks will fail to achieve a turnaround, and so it will still be necessary to have a well-designed plan of structured intervention (redefined in terms of transparent market-value triggers such as the QMVER) and a workable rapid-resolution plan. But even in this case, the regulatory authorities will have gained a significant amount of time to prepare and should be able to avoid making the costly blunders that often accompanied the hastily arranged bailouts in the crisis of 2008.

CoCos have an additional value when we are unsure about the optimal amount of equity capital that a bank should issue. Banks, like other firms, are subject to numerous frictions that would lead to the choice of an optimal equity-to-asset ratio—perhaps one similar to that chosen by finance companies that do not benefit from access to the safety net. But banks are so thoroughly entangled in the safety net that it is very difficult to determine what ratio they would choose if no safety net were available and they were forced to internalize the costs of their mistakes. This is an interesting experiment that could be run, but no modern society has the political will to undertake it and so, inevitably, capital requirements will be somewhat arbitrary. Based on recent experience, we can be sure that banks should be required to issue more equity capital, but how much more remains unclear. If the costs of issuing equity capital were entirely negligible, it would be prudent to err on the side of caution. But if regulation becomes too costly—not just in terms of a required capital ratio that diverges from the optimum, but also in terms of the growing costs of compliance with a rapidly expanding set of complex regulations—systemically important activity may shift from banks to shadow banks. This is the unpleasant trade-off that supervisors have long faced: Is it better to keep crime on the streets where it can be monitored and controlled? Or should it be pushed into dark alleys? Although CoCos cannot solve this dilemma, they can induce banks to maintain higher equity buffers and stronger risk controls than an equity requirement alone and they permit shareholders to enjoy some of the benefits of the tax shield open to other corporations without increasing the risk of default.[11] Moreover, the issuance costs of CoCos in normal times should be less than the issuance costs of new equity in times of stress to stay above regulatory minimums, and so suitably designed CoCos may provide more flexibility to the financial system when it is most needed.

[11] Of course, it would be preferable to abolish this tax distortion completely, but as long as it exists, denying banks the opportunity to issue tax-deductible debt while permitting other firms to do so is likely to encourage the migration of a considerable amount of traditional bank activities to other, less-regulated firms. A decade ago, it might have been credible to assert that as long as the deposit function was protected, this was of no consequence. But in the recent crisis, governments felt obliged to bail out numerous firms that offered negligible, if any, deposits.

References

Admati, A., P. DeMarzo, M. Hellwig, and P. Pfleiderer. 2011. "Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity is *Not* Expensive." Working Paper, Stanford Graduate School of Business (March 23).

Alessandri, P., and A. Haldane. 2009. "Banking on the State." Presented at the 12th Annual Federal Reserve Bank of Chicago International Banking Conference, September 25. Available at http://www.bankofengland.co.uk/publications/speeches/2009/speech409.pdf.

Basel Committee on Banking Supervision. 2011. "Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems." Revision of December 2010 Proposal, Bank for International Settlements (June).

Berger, A. N., R. J. Herring, and G. P. Szegö. 1995. "The Role of Capital in Financial Institutions." *Journal of Banking and Finance* 19 (3–4): 393–430.

Calomiris, C. W., and R. J. Herring. 2011. "Why and How to Design a Contingent Convertible Debt Requirement." Working Paper, Wharton Financial Institutions Center.

Federal Reserve Board. 2011. "Flow of Funds Accounts of the United States." Z1 Statistical Release, Flow of Funds Tables F110 and F121. *Federal Reserve Bulletin* (September 16). Available at http://www.federalreserve.gov/releases/z1/Current/z1r-3.pdf.

Flannery, M. J. 2005. "No Pain, No Gain? Effecting Market Discipline via 'Reverse Convertible Debentures'." In *Capital Adequacy Beyond Basel: Banking, Securities, and Insurance*, ed. H. S. Scott (chapter 5). Oxford: Oxford University Press.

Herring, R. J. Forthcoming. "Fair Value Accounting, Disclosure and Financial Stability: Does How We Keep Score Influence How the Game is Played?" In *Encyclopedia of Financial Globalization*, ed. G. Caprio. North-Holland Press.

Institute of International Finance. 2010. "Interim Report on the Cumulative Impact on the Global Economy of Proposed Changes in the Banking Regulatory Framework."

Miles, D., J. Yang, and G. Marcheggiano. 2011. "Optimal Bank Capital." Bank of England Discussion Paper No. 31 (April).

Miller, M. H. 1995. "Do the M&M Propositions Apply to Banks?" *Journal of Banking and Finance* 19 (3–4): 483–89.

Modigliani, F., and M. H. Miller. 1958. "The Cost of Capital, Corporation Finance, and the Theory of Investment." *American Economic Review* 48 (3): 261–97.

Schuermann, T. 2011. "How Much Capital for Banks? Reaction to Admati et al." Remarks delivered at the Western Economic Association International 86th Annual Conference, San Diego, California, June 30.

Tarullo, D. K. 2011. "Regulating Systemically Important Financial Firms." Speech presented at the Peter G. Peterson Institute for International Economics, Washington, DC, June 3.

Wall Street Journal. 2011. "Tarullo's Capital Idea: A Better Way to Build Safer Banks than Dodd-Frank's Regulations." Review & Outlook section, June 16.

Anchoring Countercyclical Capital Buffers: The Role of Credit Aggregates*

Mathias Drehmann, Claudio Borio, and
Kostas Tsatsaronis
Monetary and Economic Department,
Bank for International Settlements

We investigate the performance of different variables as anchors for setting the level of the countercyclical regulatory capital buffer requirements for banks. The gap between the ratio of credit to GDP and its long-term backward-looking trend performs best as an indicator for the accumulation of capital, because this variable captures the build-up of systemwide vulnerabilities that typically lead to banking crises. Other indicators, such as credit spreads, are better at indicating the release phase, as they are contemporaneous signals of banking sector distress that can precede a credit crunch.

JEL Codes: E44, E61, G21.

1. Introduction

Financial boom-and-bust cycles are costly for the banks involved and for the economy at large. Between mid-2007 and end-2010, major global banking institutions reported cumulative write-downs to the tune of $1.3 trillion. Output declined dramatically. The cumulative impact over 2008–10 on economic activity in the harder-hit

 We would like to thank Òscar Jordà for very constructive comments as well as Leonardo Gambacorta, Gabriel Jiménez, Rafael Repullo, Jesùs Saurina, Carlos Trucharte, and Mark Zelmer for discussions and comments on previous drafts, and Jakob Demski for much appreciated work with the data collection. The views expressed in this paper are of the authors' own and do not necessarily reflect those of the Bank for International Settlements or the Basel Committee on Banking Supervision. Corresponding authors: Mathias Drehmann: mathias.drehmann@bis.org (tel: +41 61 280 8089); Claudio Borio: claudio.borio@bis.org (tel: +41 61 280 8436); Kostas Tsatsaronis: ktsatsaronis@bis.org (tel: +41 61 280 8082).

advanced economies exceeded 10 percent of their respective GDP, and average unemployment rates shot up from about 5 percent to nearly 9 percent. The repercussions of the crisis were felt by countries outside its epicenter. Between mid-2008 and mid-2009, world GDP contracted by 1.6 percent for the first time in recent memory (IMF World Economic Outlook 2011). Unsurprisingly, the experience added impetus to policymakers' and academic economists' efforts to better understand the mechanisms that drive financial system procyclicality and to devise policy tools that can mitigate it.

This paper examines one such tool: time-varying regulatory capital buffers for banks. It focuses specifically on the choice of indicators that can provide a reliable guide for regulatory capital requirements to dampen banks' procyclical behavior, restraining risk taking during booms and cushioning financial distress during busts. We analyze the behavior of a wide range of possible indicator variables around episodes of systemic banking crises, drawing on the empirical evidence from more than forty crises in thirty-six countries.

The analysis focuses separately on the run-up phase to the crisis and on the phase that follows its outbreak. This is necessary because financial stability risks tend to build up gradually in good times, but their consequences materialize quite suddenly.[1] This also means that the requirements for the policy tool differ between these two phases. Early-warning properties are very important in the phase in which vulnerabilities build up, so as to activate policy tools in time to influence behavior. By contrast, the ability to signal banking sector stress in real time is critical in guiding the tool during a crisis.

We find that the variable that performs best as an indicator for the build-up phase is the gap between the ratio of credit to GDP and its long-term trend (the credit-to-GDP gap). Across countries and crisis episodes, the variable exhibits very good signaling properties, as rapid credit growth lifts the gap as early as three or four years prior to the crisis, allowing banks to build up capital with sufficient lead time. In addition, the gap typically generates very low "noise," by not producing many false warning signals that crises are imminent.

The credit-to-GDP gap, however, is not a reliable coincident indicator of systemic stress in the banking sector. In general, a prompt and sizable release of the buffer is desirable. Banks would then be free

[1] Jiménez and Saurina (2006) provide empirical evidence for Spain.

to use the capital to absorb write-downs. A gradual release would reduce the buffer's effectiveness. Aggregate credit often grows even as strains materialize in the banking system. This reflects in part borrowers' ability to draw on existing credit lines and banks' reluctance to call loans as they tighten standards on new ones. A fall in GDP can also push the ratio higher. Aggregate credit spreads do a better job in signaling stress. However, their signal is very noisy: all too often they would have called for a release of capital at the wrong time. Moreover, as spread data do not exist for a number of countries, their applicability would be highly constrained internationally.

We conclude that it would be difficult for a policy tool to rely on a single indicator as a guide across all cyclical phases. It could be possible to construct rules based on a range of conditioning variables rather than just one, something not analyzed in this paper. However, it is hard to envisage how this could be done in a simple, robust, and transparent way. More generally, our analysis shows that all indicators provide false signals. Thus, no fully rule-based mechanism is perfect. Some degree of judgment, both for the build-up and particularly for the release phase, would be inevitable when setting countercyclical capital buffers in practice. That said, the analysis of the political economy of how judgment can be incorporated in a way that preserves transparency and accountability of the policymakers in charge goes beyond the scope of this paper.

While the discussion in the paper is exclusively in terms of the design of a countercyclical buffer tool, the analysis applies to any time-varying instrument aimed at reducing procyclicality that relies on indicator variables. The behavior of different indicator variables in the build-up and release phases is the key parameter determining their suitability.

The rest of the paper is organized as follows. Section 2 frames the issues by discussing the objectives of the countercyclical capital buffer and placing this work in the broader context of the literature. Section 3 discusses the desirable characteristics of an indicator variable. Section 4 describes the candidate variables we analyze and explains the data used in constructing them. Section 5 explains the statistical exercises, conducted separately for the build-up and release phases. Section 6 presents the results of some robustness analysis, concerning the choice of detrending parameters and dealing with the cross-country exposures of banks. The last section concludes.

2. The Main Objective of Countercyclical Capital Buffers

Financial system procyclicality refers to the mutually reinforcing interactions between the real and financial sectors of the economy that tend to amplify the business cycle and that are often at the root of financial instability.[2] These interactions are most evident during periods of financial stress. A weakened financial system facing strains cannot absorb further losses without retrenching from risk and credit intermediation, leading to fire sales and a “credit crunch.”[3] Arguably, however, the seeds of the strains that materialize during downturns are sowed earlier, during the preceding upswing. Episodes of severe financial sector stress are typically preceded by extended periods of unusually low perceived risk, marked by booming financial sector activity and strong asset price growth (e.g., Reinhart and Rogoff 2009). Amplifying feedback mechanisms can be as potent in the expansion phase as they are in cyclical downturns.

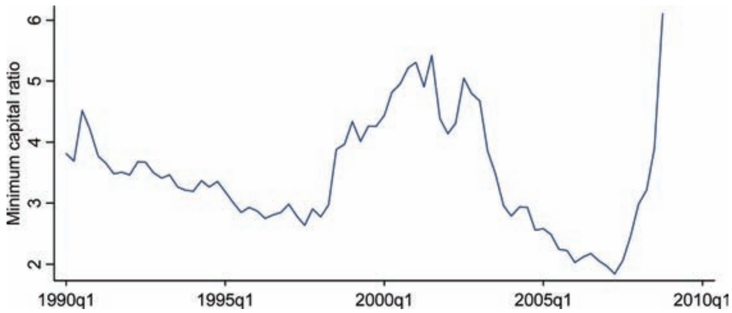
There is a long tradition, most prominently expressed by Minsky (1982) and Kindleberger (2000), which sees financial crises as the result of mutually reinforcing processes between the financial and real sides of the economy. In this view, financial imbalances are driven by, but also feed, an unsustainable economic expansion, which manifests itself in unusually rapid growth of credit and asset prices. As the economy grows, cash flows, incomes, and asset prices rise; risk appetite increases; and external funding constraints weaken. This, in turn, facilitates risk taking. The financial system typically does not build up sufficient capital and liquidity buffers during benign economic conditions, when it is easier and cheaper to do so, in order to deal with more challenging times. At some point, imbalances have to unwind, potentially causing a crisis, characterized by large losses, liquidity squeezes, and possibly a credit crunch.

Financial system procyclicality can be traced to two fundamental sources (e.g., Borio 2003, Bank for International Settlements 2009). The first source is limitations in risk measurement. The typical

[2]For a fuller discussion, see Borio, Furfine, and Lowe (2001), Brunnermeier et al. (2009), and references therein.

[3]For evidence related to the recent crisis, see Ivashina and Scharfstein (2010).

Figure 1. Procyclical Assessment of Credit Risk



Notes: The capital charge (in percentage points) is a measure of portfolio credit risk based on the risk weights embedded in the internal ratings-based (IRB) methodology of Basel II. The IRB formula is applied to the exposures in a credit portfolio that includes rating categories ranging from AAA to Caa3. We follow Catarineu-Rabell, Jackson, and Tsomocos (2005) in setting the portfolio shares of each rating class. For each exposure and observation in the sample, we proxy the probability of default with the expected default frequency taken from the database of Moody's KMV.

parameters underlying risk-measurement practices tend to be procyclical (e.g., Gordy and Howells 2006). The second, complementary, source is distortions in incentives (e.g., Rajan 2005). Principal-agent issues result in conflicts of interest between providers and users of funds. In addition, externalities in behavior and strategic complementarities suggest that actions that may be rational from the perspective of individual agents may collectively result in undesirable outcomes. Short horizons by private decisionmakers enhance the impact of both fundamental sources on procyclicality.

Figure 1 illustrates the cyclicality in risk measures, using as a metric the capital charge that would apply on a hypothetical credit portfolio. The calculation applies the risk weights in the internal ratings-based (IRB) methodology of Basel II and uses market-derived probabilities of default as the main input. The cyclicality of the risk measure is clearly apparent. Focusing on the most recent period, capital would have reached its all-time minimum just before the crisis in 2007, after which it would have increased rapidly.

Against this backdrop, what should be the objective of countercyclical policy tools and, in our specific case, of countercyclical bank capital requirements?

We distinguish three objectives that vary in their degree of ambition for countercyclical policies. The first, most ambitious, objective is to smooth the *business cycle* through the influence of capital requirements on banks, i.e., to use them as a demand management tool. Setting such an objective is commensurate to calibrating time-varying prudential capital requirements to achieve a *macroeconomic goal*. A less ambitious objective is to smooth the financial (credit) cycle. We think of this approach as using capital requirements as a means of achieving a *broad macroprudential goal*. The third, and least ambitious, objective is simply to protect the banks from the build-up of systemwide vulnerabilities. We call this a *narrow macroprudential goal*, designed to strengthen systemic resilience without taking explicit account of its influence on the financial and business cycles.

The choice of objective is not independent of the nature of the instrument. We argue that it would be prudent for the objective of the countercyclical buffer to be relatively modest. This argument is based both on the effect of capital requirements on economic behavior and on the current state of our knowledge of the quantitative links between capital requirements, credit availability, and economic activity.

A macroeconomic stabilization objective would make the analysis in this paper superfluous: a natural candidate to anchor the instrument would be a standard measure of economic activity, such as GDP growth. At the same time, however, choosing this objective would overestimate the current state of our knowledge about the influence of capital requirements on the credit cycle and output. The literature on dynamic models that combine financial and real sectors is growing (e.g., Bernanke, Gertler, and Gilchrist 1996, Van den Heuvel 2008, Gerali et al. 2010, and Meh and Moran 2010).[4] However, the models are still very much in their infancy and do not yet match the level of maturity and general acceptance that characterize the prevailing paradigm in the field of monetary policy.[5] Moreover, our empirical understanding of the impact of changes in

[4]Other examples that build on the financial accelerator literature developed by Bernanke and Gertler (1989) and the credit cycles model of Kiyotaki and Moore (1997) include Lorenzoni (2008) and Korinek (2011). Zhu (2008) analyzes a dynamic model linking bank capital to credit and output.

[5]A representative example of the prevailing paradigm for the analysis of monetary policy is Woodford (2003).

regulatory instruments (such as capital requirements) on lending, asset prices, and, by extension, aggregate expenditure and GDP is still very limited. To be sure, the literature on the credit channel has largely supported the notion that higher capital requirements would have a negative influence on the supply of credit, at least for institutions that are weakly capitalized and illiquid (for surveys, see VanHoose 2007 and, more recently, Gambacorta and Marques-Ibanez 2011).[6] But it is rather uninformative about whether they would have a similar effect during a boom. Also, the link between credit and GDP tends to be time varying and dependent on the financial structure of the economy.[7]

Similar arguments would apply to adopting the broad macroprudential objective of smoothing the credit cycle. In this case, the relevant objections would concern only the link between capital requirements, on the one hand, and credit supply and bank risk taking, on the other, and not the impact of the requirements on aggregate demand and GDP. At the same time, however, the design of the instrument would at least require an operational definition of the credit (or financial) cycle.

The literature on the credit cycle and its relationship to the business cycle provides little guidance. It relies mostly on simple setups in which the credit cycle is *fully aligned* with the business cycle. For example, in Kashyap and Stein (2004) a social planner wants not only to protect the economy from externalities arising from possible bank defaults but also to ensure that positive net-present-value projects are funded. However, in states where banks experience high loss rates and raising capital is very costly, the supply of credit is constrained. In these cases, the optimal balance for a policymaker is to accept higher failure rates. Repullo and Suarez (2009) identify a similar trade-off.

Importantly, in these examples good and bad states are identified respectively with GDP growth and recession periods. Consequently, *by construction*, the best indicator variable to anchor countercyclical

[6]Taking the results of Gambacorta and Marques-Ibanez (2011) at face value, Drehmann and Gambacorta (2011) show that countercyclical capital buffers would have significantly reduced bank lending in Spain.

[7]For a discussion in the euro-area context, see Angelini, Kashyap, and Mojon (2003).

Table 1. Real Credit Growth during Recessions (in percent)

	Mean	Percentile							# Obs.
		5%	10%	25%	50%	75%	90%	95%	
All Data	0.3	−10.7	−5.4	−1.9	1.0	3.7	6.2	8.8	156
Until 2007:Q2	−0.1	−11.8	−6.8	−2.1	0.3	3.6	6.3	8.9	124
Current Crisis	1.9	−4.1	−1.6	−0.2	1.6	4.1	5.4	7.8	32

Note: The table shows the distribution of average real credit growth during recessions, defined as periods of at least two consecutive quarters of negative real GDP growth.

capital buffers is a measure of the business cycle, such as GDP growth. This is true regardless of whether one adopts a broad or a narrow objective. Repullo and Saurina (2011) argue that GDP growth should be the guide for countercyclical capital requirements using essentially this rationale.

Empirically, however, the business and credit cycles do not coincide. For instance, Koopman and Lucas (2005) show that at typical business-cycle frequencies, around four to eight years, there is no cyclical co-movement between GDP and default rates, although some correlation is evident at longer horizons, of eleven to sixteen years. In addition, Mendoza and Terrones (2008) find that output increases during credit booms but that output booms need not involve credit booms, which tends to reduce the degree of co-movement of the two cycles. Finally, Aikman, Haldane, and Nelson (2010) as well as Claessens, Kose, and Terrones (2011) conclude that financial cycles are longer and more pronounced than business cycles.

Table 1 illustrates how business and credit cycles are not fully synchronous. Over the period covered by our data (see next section), the average correlation between the growth rates of real credit and real GDP across the forty countries we study is about 44 percent. This is significantly different from zero, but it hardly suggests a close alignment of the two cycles. Furthermore, and more to the point for the focus of this paper, not every recession is characterized by serious credit constraints. As a rough measure, table 1 shows that for more than half of the recessions in our sample, real credit growth is actually positive.

Hence, in this paper we adopt the third and least ambitious objective, which focuses exclusively on protecting the banks from the build-up of systemwide vulnerabilities. The countercyclical capital buffer tries to accomplish this by actively encouraging the build-up of buffers in boom times (when risks are taken on but, arguably, are not fully reflected in prices) and by releasing them in bad times (when the market price of risk shoots up once losses materialize). Clearly, to the extent that a policy instrument succeeds in this narrower objective, it is also likely to make a contribution to the broader goal of smoothing the financial and business cycles. This, however, is seen as a collateral benefit rather than as the principal objective.

3. Key Characteristics of an Effective Instrument

What are the criteria that should guide the choice of the anchor variable? Given the objective of strengthening the defenses of banks against systemic risk, the criteria for the indicator variable follow from the desirable features of the countercyclical buffer.

The main idea of a countercyclical buffer is to promote the build-up of sufficient capital cushions in the banking system during the boom phase of the financial cycle and to encourage their use during stressful periods, thereby easing the strains in credit supply. From this perspective, the instrument should be designed to meet four criteria:

(i) It should signal the proper *timing* for the accumulation and release of the capital buffer. This means that it should identify good and bad times.

(ii) It should ensure that the *size* of the buffer built up in good times is sufficient to absorb subsequent losses, when these materialize, without triggering serious strains.

(iii) It should be *robust to regulatory arbitrage*. This includes being difficult to manipulate by individual institutions as well as being applicable to banking organizations that operate across borders.

(iv) It should be as *rule based as possible*, *transparent*, and *cost-effective.*

The first criterion relates to the all-important issue of characterizing the cycle against which the instrument should lean (act countercyclically). This is the focus of the empirical analysis in this paper. It is key, therefore, to characterize what we mean by "good times," when the capital buffers need to be accumulated, and by "bad times," when they should be used to absorb losses.

Kashyap and Stein (2004) and Repullo and Suarez (2009) argue that bad times are periods when banks experience high losses and the banking sector is a source of credit constraints, which in their setup coincides with GDP declines. This suggests that bad times can be identified by a mix of two factors: some measure of banks' aggregate gross losses and of the extent to which banks are a source of credit tightening. The transition from bad to good times could be identified in a similar way, but its precise timing is less critical. This is because of the asymmetry in the financial cycle. The emergence of financial strains tends to be very abrupt and, typically, comes as a surprise. It is therefore essential that the buffer is released sufficiently promptly and in sufficient amounts. By contrast, the transition from bad to good times is much more gradual.

Finding good measures for losses and credit conditions is often problematic. Aggregate loss series are not widely available and accounting rules tend to distort their timing. In practice, loan-loss provisions tend to behave as lagging rather than contemporaneous indicators of bank distress. Credit conditions are measured in several countries by surveys, such as the Loan Officer Opinion Survey in the United States. These surveys relate to changes in credit conditions, not to the absolute degree of tightness. By construction, therefore, they can point to an easing of conditions even as credit supply is severely constrained.[8] In addition, survey-based measures could be subject to strategic reporting were they to be used to anchor countercyclical capital requirements. Finally, they are not widely available internationally.

Instead of relying on banking sector losses combined with a measure of credit conditions, we use historical banking crises as empirical

[8]Even though it measures only the change in credit conditions, the net-tightening series in the United States were found to be very helpful in anticipating a credit crunch and its effect on the business cycle (Lown, Morgan, and Rohatgi 2000 and Lown and Morgan 2006).

proxies for bad times.[9] The key benefit of this approach is that data on historical banking crises are widely available for a large set of countries going back in time. Given the identification of bad times with banking crises, our empirical strategy is to find indicators which would lead to a build-up of capital buffers ahead of crises, i.e., during the good times. Equally, we assess whether there are variables which signal a release of capital buffers at the onset of banking crises.

The second criterion implies that the variation in the indicator variable should be sufficient to provide a meaningful quantitative guide for the accumulation and release phases. In particular, the signals should be comparable across time and noise-free, avoiding unnecessary reversals of direction from one period to the next. Empirically, this criterion rules out bank-specific indicators. Because of idiosyncratic factors, these tend to fluctuate widely from one year to the next, so that buffers would be built up and released in short succession (Drehmann et al. 2010). Such volatility would wreak havoc in banks' capital planning and would likely encourage banks to treat the countercyclical buffer as the new minimum. Therefore, we do not discuss bank-specific variables in what follows.

The third criterion is self-evident. To the extent possible, regulatory arbitrage should be minimized both within and across borders. And since finance is global, the design should take into account the fact that banks are typically exposed to financial cycles in multiple jurisdictions.

The fourth criterion covers a range of aspects. Rules are especially appealing because of the political economy obstacles that hinder the build-up of buffers during booms. Transparency is needed to support appropriate governance, particularly if strict rules are not feasible and some judgment is required. Cost-effectiveness favors continuity and seamless integration with the rest of the regulatory framework. It suggests that it would be helpful to express the buffer in terms of risk-weighted assets and as an add-on to the regulatory minimum level of capital.[10] Importantly, the scheme would thus

[9]Drehmann et al. (2010) show that for the United States, for which most of the relevant data are available, banking crises are the only periods when both banking sector losses are high and credit conditions are tightened.

[10]This requirement prevents adjustments that lower the minimum in bad times, as suggested, for example, by Gordy (2009).

retain the cross-sectional differentiation of risk at a given point in time while counterbalancing the tendency of most widely used risk measures to vary procyclically (i.e., to assess risk as low in good times and high in bad times, as illustrated in figure 1).

To ensure robustness and transparency, we considered a representative set of *single* indicator variables as possible anchors for the buffer. As will be shown, single indicators already provide very good guidance, leaving limited scope for incremental improvement through the use of multivariate approaches.[11] Arguably, no rule-based method can fully capture the complex dynamics of financial cycles. Some degree of judgment will always be required.

4. Different Candidates for Anchor Variable

As mentioned above, the anchor variable is best viewed as a proxy for the underlying cyclicality addressed by the instrument. We therefore classify the variables in three categories that correspond to different aspects of the financial cycle: the macroeconomy, banking sector activity, and funding costs. In this section we briefly discuss the pros and cons of these variables, the data used to construct them, and their behavior around episodes of systemic stress.

4.1 The Macroeconomy

Variables that relate to the macroeconomy capture broad trends in the financial and real sectors; as such, they are rough summary measures of aspects of the financial cycle. They also have the advantage of being immune to strategic manipulation by individual institutions. We assess the indicator properties for a number of variables corresponding to real economic activity, financial quantities, and asset prices.[12] These variables are, of course, influenced by the collective

[11] Borio and Drehmann (2009a) and Borio and Lowe (2002) show that combinations of variables have somewhat better signaling properties for systemic financial distress than single indicators.

[12] We also assessed inflation. However, the theoretical link between inflation and systemic risk is unclear and, given its very weak performance, we do not report the results for the sake of brevity.

behavior of banks, but in a reasonably competitive market any single institution would view them as exogenous. In addition, most macroeconomic series are widely available and therefore could be used in many countries.

Real GDP. We consider annual real GDP growth and the (real) output gap. These are the most natural indicators of the aggregate business cycle. That said, as already discussed, the business and the financial cycles, although closely linked, are not fully synchronized.

Real Credit. The cycle is often defined with reference to credit availability. Aggregate real credit growth (annual) could be a natural measure of the credit cycle—in particular, if not only bank credit but all other sources of credit are taken into account. As credit to the private sector tends to grow rapidly during booms and slow down or contract during credit crunches, deviations of credit growth from a trend could be an informative variable. Due to data limitations, we focus on bank credit to the private non-financial sector in our analysis except for the United States, for which we use a broad credit measure.

We exclude public-sector debt from our analysis, as it is countercyclical. It tends to slow down in booms and rise rapidly after stress materializes. Data availability is also an issue, as for many countries information is only available annually. Using annual data for a subset of our sample, we found that the inclusion of public-sector debt severely reduces the performance of credit-related variables, in that they indicate fewer crises and issue more false signals. For brevity, these results are not reported but are available on request.

Credit Relative to GDP. Here we consider two related indicators. The first is the difference between the annual growth of credit (to the private non-financial sector) and the annual growth of output, and the second is the credit-to-GDP gap. Both indicators benchmark credit growth on the growth of overall economic activity, trying to capture whether credit is booming or contracting "excessively" relative to GDP. The difference between the two growth rates performs this comparison at the business-cycle frequency, assuming a constant long-term trend in the credit-to-GDP relationship. By contrast, deviations of the credit-to-GDP ratio from its long-term trend (the "credit-to-GDP gap") are more sensitive to lower-frequency structural changes, such as natural financial deepening.

Monetary Aggregates. In the simplest macro models, money and credit are virtually interchangeable indicators, being the two sides of a simplified bank's balance sheet. However, in both theory and practice, the two do not coincide. Behaviorally, their links with asset prices and asset returns, in particular, are very different. And it is known that the credit-to-deposit ratio has a marked cyclical pattern, notably rising during booms. Banks can fund themselves through non-monetary sources (e.g., wholesale interbank funding) and shift their assets between government securities and credit to the private sector.[13] Empirically, it has also been shown that credit and monetary aggregates series decoupled after the Second World War (see, for example, Schularick and Taylor, forthcoming). Thus, real monetary growth, in our case measured by the annual growth rate in M2, may provide an alternative measure of the financial cycle.

Asset Prices. Asset prices in general, and property prices in particular, tend to show exceptionally strong growth ahead of systemic banking events. They also fall precipitously during periods of financial stress. We therefore consider the annual (real) growth rate of equity prices and property prices. Property prices are a weighted average of residential and commercial property prices, where weights are based on estimates of the relative market shares in each country. We also consider deviations from long-term trends, as equity and property price gaps have proved useful in predicting banking crises (e.g., Borio and Drehmann 2009a).

All gaps are calculated as differences from a *one-sided* Hodrick-Prescott filter. This way the calculation of the trend considers only information that would have been available at the time the buffer is activated, as it excludes the path of the given variable at future dates. The specification of the filter is discussed in the data subsection below.

4.2 Banking Sector Activity

Aggregate measures of bank activity tend to co-move with the business and financial cycles. During periods of high bank profitability, banks tend to increase their intermediation activity through

[13]See Borio and Lowe (2004) for a theoretical and empirical analysis of this issue, including an examination of the comparative leading indicator properties of the two variables.

rapid credit growth and to take on risks. Benign economic conditions are associated with low credit losses and high internal capital resources (retained earnings), as well as cheap and easily available external funding. As a result, the cost of accumulating buffers is comparatively low.

Banking Sector Profits. This is a key indicator of the sector's performance. Earnings are high in good times and reflect losses in times of stress. Admittedly, profit figures can be subject to strategic management by banks, something that may distort their information content. That said, the scope is partly constrained by the scrutiny of analysts, shareholders, and regulators.

Aggregate Gross Losses. This indicator of performance focuses on the cost side (non-performing loans, provisions, etc). The financial cycle is frequently signaled by the fall and rise of realized losses.

4.3 Cost of Funding

This category focuses on the cost to banks of raising funds. By identifying the cycle with fluctuations in the cost of funding, a regulatory rule would incentivize banks to raise funds when these are relatively cheap and allow them to use the buffers in periods of stress, when such funding becomes more expensive.

Banking Sector Credit Spreads (Indices). These are indicators of vulnerabilities in the banking sector, reflecting markets' assessment of the risk of bank failures. By being closely tied to the financial condition of banks, they may be subject to manipulation. Relying on broad indices, where they exist, can mitigate this drawback. In the analysis we consider the average of credit default swaps (CDS) spreads for the largest banks in each country.[14]

Cost of Liquidity. These are indicators of the banking sector's average cost of raising short-term funds. They are closely linked to banks' health and aggregate funding conditions in markets. In normal times interbank markets distribute liquidity seemingly without friction. When severe strains emerge, measures of funding costs,

[14]Gordy (2009) argues in favor of CDS spreads as an anchor variable for a countercyclical buffer.

such as the LIBOR rate, tend to jump. These indicators may therefore be ideal in marking the transition from good to bad times. However, many interbank market rates may be unrepresentative of actual funding conditions. In a crisis, the dispersion in credit quality across banks tends to increase and institutions have a greater incentive to strategically misreport their borrowing costs (Gyntelberg and Wooldridge 2008). Interbank rates, such as LIBOR, which are not based on actual transactions but are the outcome of a survey amongst a panel of banks, could be subject to strategic manipulation. In the analysis we consider three-month LIBOR-OIS spreads, i.e., the difference between the three-month interbank rate minus the rate in three-month overnight index swaps.[15]

Corporate Bond Spreads (Average). This is an indicator of credit quality for the economy at large. During boom phases, spreads are typically lower than average, while they tend to widen suddenly and sharply during periods of stress. Spreads can also be viewed as indicators of the average cost of borrowing in the economy, including by banks. They can thus be used as an anchor for a policy tool that seeks to smooth funding costs. In this analysis, we consider the spread between the yield on BBB corporate bonds and government bonds.

4.4 Data

We analyze thirty-six countries (plus the euro area for some market indicators).[16] The period of analysis starts in 1960 for some countries and series, and at the earliest available date for the rest. All data are quarterly, except for aggregate profits and losses, which are annual.

[15]The spread between government paper and the Eurodollar deposit rate (the so-called TED spread) provides very similar information. For the sake of brevity, these results are not shown here.

[16]Drehmann, Borio, and Tsatsaronis (2011) provide a detailed overview of the data in the sample. The countries included in the analysis are Argentina, Australia, Austria, Belgium, Canada, Chile, China, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong, India, Indonesia, Ireland, Italy, Japan, Korea, Lithuania, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Portugal, Singapore, South Africa, Spain, Sweden, Switzerland, Taiwan, Turkey, the United Kingdom, and the United States.

The performance of the anchor variables for the countercyclical capital buffer is assessed against an indicator of banking crises. Admittedly, the dating of banking crises is not uncontroversial (e.g., Boyd, de Nicoló, and Loukianova 2009). We follow the dating of crises in Laeven and Valencia (2008, 2010) as well as in Reinhart and Rogoff (2009). In addition, we use judgment and draw on correspondence with central banks to determine some of the crisis dates. This results in forty-nine different crises. The full list of crisis dates is given in Drehmann, Borio, and Tsatsaronis (2011).

Macroeconomic variables are generally available for all countries and are collected from national authorities, the International Monetary Fund international financial statistics, and the Bank for International Settlements (BIS) database. Property prices are based on BIS statistics and are available only for eighteen countries.

When a variable is expressed as a gap (i.e., as the difference between the current level and the long-term trend),[17] we measure the trend with a one-sided Hodrick-Prescott filter. The backward-looking filter is run recursively for each period and the gap calculated as the difference between the actual value of the variable and the value of the trend at that point. Thus, a GDP trend calculated in, say, 1988:Q1 only takes into account information up to 1988:Q1, and the GDP trend in 2008:Q4 takes into account all information up to 2008:Q4. This is an important practical constraint, as policymakers have to take decisions in real time and rely on data that are available at that point. Before using any trend, we require at least five years of information.[18]

The calculation of the Hodrick-Prescott filter involves a key smoothing parameter λ. Following Hodrick and Prescott (1981), it has become standard to set the smoothing parameter λ to 1,600 for quarterly data. Ravn and Uhlig (2002) show that for series of other frequencies (daily, annual, etc.), it is optimal to set λ equal to

[17]For asset price gaps, the difference between the actual data and the trend at each point in time is normalized by the trend in that period. For the credit-to-GDP gap, we simply take the difference between the actual data and the trend at each point in time.

[18]Ideally, ten years of data would be better (e.g., Borio and Lowe 2002). But given that data are limited for some series in some countries, we chose a five-year window to ensure sufficient observations.

1,600 multiplied by the fourth power of the observation frequency ratio. We set λ for *all* the gaps to 400,000, implying that financial cycles are four times longer than standard business cycles.[19] This seems appropriate, as crises occur on average once in twenty to twenty-five years in our sample. Thus in the main section, λ is set to 400,000 to derive the output, the credit-to-GDP, the property, and the equity gaps. For robustness, we analyzed alternative values of the smoothing parameter and discuss the results in section 6.1.

Data on aggregate profits for banks are hard to obtain. We rely on the OECD banking statistics.[20] Specifically, we use aggregate net provisions, as an indicator of gross losses, and profits before tax. Both are normalized by total assets and are only available on an annual basis. For practical purposes, annual data are unlikely to be sufficient—in particular, when considering the release of the buffer. But the OECD database is the only source for most countries.

Given the heterogeneity in data availability, our analysis for macro and banking sector conditions considers two data sets. The first data set includes all available data and thus uses a different period for the analysis of each variable. The second is a homogenous data set that includes only observations for which all macro variables (including property prices) as well as profit and gross loss indicators are available.

A full analysis of market-based indicators is impossible. Most relevant data start only in the late 1990s and are only available for few countries. Only four crisis episodes fall within the corresponding sample, and three of them are in 2007. Therefore, the performance covers only the most recent crises. For the sake of completeness, we report the analysis on this sample as well. But we do not want to emphasize the results, as they cannot be statistically robust.

[19]This is the same value as in previous comparable work (e.g., Borio and Lowe 2004 and Borio and Drehmann 2009b).

[20]While the OECD data are consistent across countries and broadly available, for some countries (in particular, for the United States and the United Kingdom) they indicate lower profits and losses than other national sources. However, the correlation between different data series in each country is high and typically well above 75 percent. Other data sources are used by Drehmann et al. (2010) for a subset of six countries and yield the same qualitative results as those presented here.

4.5 The Behavior of Candidate Variables around Systemic Crises

As a first step, we look at the performance of different indicator variables around episodes of systemic banking crises. Figures 2–4 summarize the behavior of each variable during a window of sixteen quarters before and after the crisis date (time 0 in the figures). For each variable we use data from as many countries as possible and show the median (solid line) as well as the 25th and 75th percentiles (dashed lines) of the distribution across episodes. The figures provide some insight into how different indicator variables behave during the accumulation and release phases of the capital buffer.

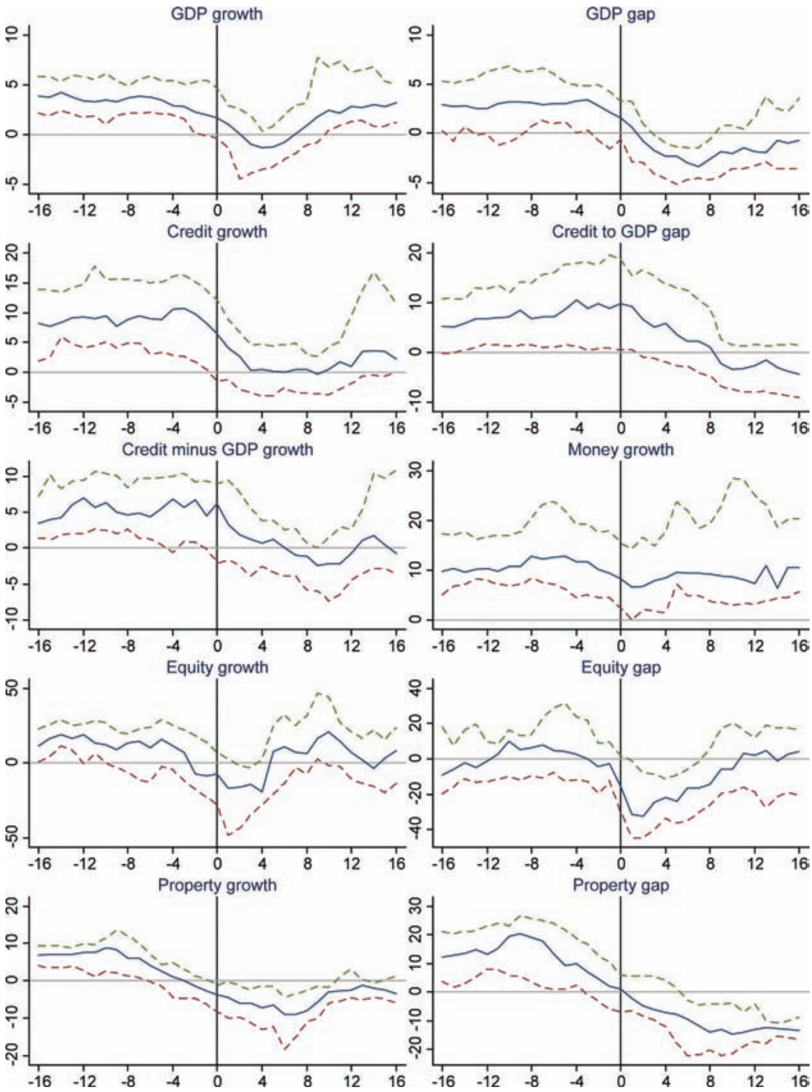
Median real GDP growth is around 4 percent four years prior to a crisis. It then starts to decline, with a slowdown gathering momentum in the year leading up to the crisis. Once the crisis materializes, GDP growth turns negative. After around two years, on average, the economy returns to its pre-crisis growth rate, suggesting that this growth rate is not particularly unusual. Interestingly, the 75th percentile shows that many crises are not preceded by any slowdown in output.[21] In line with real GDP growth, the output gap shows a similar pattern.

Real credit growth, the difference between credit and output growth, and the credit-to-GDP gap, all rise in the lead-up to banking crises. Therefore, they could be useful indicators during the accumulation phase. For the release phase, real credit growth could provide useful information, as it falls significantly around the event. The indicators based on credit in relation to GDP, on the other hand, remain elevated for around one to two years after the crisis. Money growth shows a pattern similar to that of credit growth, even though the rise before and fall after the crisis is less pronounced. Similar to output growth, money growth quickly returns to pre-crisis levels, suggesting that these levels may not be unusual.

As expected, asset, and in particular property, prices tend to grow rapidly ahead of banking crises. This could make them useful indicators for the accumulation phase of the buffer. However, they

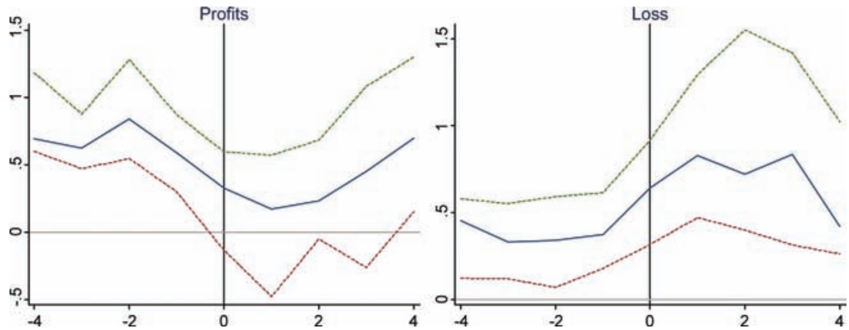
[21] This provides a clear indication that models linking fragilities in the banking sector to weak macroeconomic fundamentals, such as macro stress tests, do not capture the dynamics of many crises (Alfaro and Drehmann 2009).

Figure 2. Macroeconomic Variables around Crises



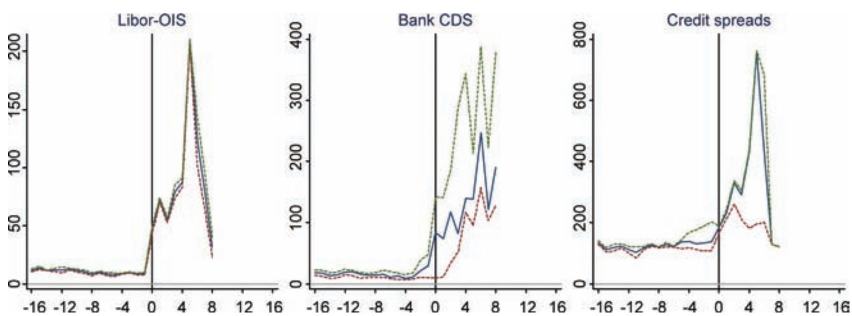
Notes: The horizontal axis depicts plus/minus sixteen quarters around a crisis, which is indicated by the vertical line. The solid line corresponds to the median across all observations in a particular quarter. The upper and lower dashed lines represent the 75th and 25th percentile, respectively.

Figure 3. Banking Sector Conditions around Crises



Notes: The horizontal axis depicts plus/minus four years around a crisis, which is indicated by the vertical line. The solid line corresponds to the median across all observations in a particular quarter before or after the onset of a crisis. The upper and lower dotted lines represent the 75th and 25th percentile, respectively.

Figure 4. Market Indicators around Crises



Notes: The horizontal axis depicts plus/minus sixteen quarters around a crisis, which is indicated by the vertical line. The solid line corresponds to the median across all observations in a particular quarter before or after the onset of a crisis. The upper and lower dotted lines represent the 75th and 25th percentile, respectively. The analysis of market variables ends in 2009:Q4, meaning that only eight quarters of data are available for these series after the beginning of the recent crisis.

tend to fall well before the onset of the crisis. If indicator variables are linked mechanically to capital requirements, this would result in a premature draw-down of the buffers.

Banks' losses increase significantly after crises. Equally, profits drop. However, there is not much variability in the run-up to a crisis, raising questions about the ability of these variables to capture the intensity of the cycle. By contrast, the variables may be useful indicators for the release phase, even though for this purpose the annual frequency is a major drawback.

Market indicators seem to perform exceptionally well as signals for the release of capital. All of them rise significantly around crisis dates. However, before a crisis they seem to be low and stable, thereby not providing clear measures of the intensity of the build-up of systemic risk. That said, given the low number of crises (at most four, in the case of credit spreads) for which market data are available, we need to put a strong caveat on these conclusions.

The discussion suggests that there is possibly no single indicator variable which works equally well for the accumulation and release phases. This is not surprising: it would require a variable that is both a coincident *and* leading indicator of systemic distress—or, in the language of Borio and Drehmann (2009b), an indicator that acts as both a barometer and a thermometer of financial distress. Therefore, we analyze the accumulation and release phases separately, starting with the former.

5. Statistical Analysis

In this section, we assess the performance of the indicators more formally by using a signal-extraction method. After a brief methodological discussion, we analyze the build-up and release phases sequentially.

5.1 Methodology

Following the literature on early-warning indicators for systemic banking crises (e.g., Kaminsky and Reinhart 1999), we use a signal-extraction method to compare the performance of different variables.

We use this method for both the build-up and the release phases, but rely on slightly different specifications in each of the two cases.

Consider the build-up phase first. Let y_t be an indicator variable and $S(y_t)$ a signal which can be 0 ("off") or 1 ("on") depending on whether y is below or above a threshold value k.[22] The signal is "on" if y_t^h exceeds a critical threshold level k, i.e., $S(y_t) = 1$ if $y_t^h > k$. For each indicator variable we assess a range of thresholds (k). A signal of 1 (0) is judged to be correct if a crisis (no crisis) occurs any time within a three-year (twelve-quarter) forward horizon.[23] In the robustness section we will also analyze horizons of one and two years.

We, thus, consider a flexible horizon, as originally suggested by Borio and Lowe (2002). An alternative assumption would be to use a fixed lead-lag relationship, so that a signal at time t would be judged to be correct if a crisis materialized exactly h periods ahead. A range of different forecast horizons could be considered. However, such an approach implicitly confounds the indication that a crisis is imminent with the prediction of its exact timing. We believe this is problematic. The dynamics of banking crises differ considerably across episodes and their exact timing is, almost by definition, unpredictable. What the indicators detect is the general build-up of vulnerabilities in the financial sector that creates the conditions for a crisis. This is the essence of our approach and the reason why we use a window during which a crisis may occur rather than a specific interval of time between signal and event.

The assessment methodology distinguishes between two types of forecast errors: type 1 error, when no signal is issued and a crisis occurs, and type 2 error, when a signal is issued but no crisis occurs. Both error types are summarized by the *noise-to-signal ratio* (NS in the tables), which is the ratio of type 2 errors (as a fraction of non-crisis dates) to 1 minus the fraction of type 1 errors (as a fraction of all crisis dates). The typical trade-off between these two types of

[22]Strictly speaking, indicator variables are of two types: those that are high during boom times (such as profits or credit growth) and those that are low during booms (such as credit spreads). For ease of exposition, we discuss only the former in the text, since the discussion of the latter is symmetric.

[23]If a series starts less than three years before the first crisis date, the series is not considered for the analysis until two years after the onset of this crisis.

errors is that when a variable captures a lot of crises (low type 1 error) it tends to overpredict their number (i.e., issue false signals and exhibit a high type 2 error).

Once in a crisis, it makes no sense to predict another crisis: the indicator has already done its job. We therefore do not consider any signals in the two years after the beginning of a crisis. The two-year window is at the low end of the estimates of the average length of crises. For instance, using the time which GDP requires to recover to its pre-crisis level as a measure of the length of crises, Cecchetti, Kohler, and Upper (2009) find that on average crises last nearly three years. In our sample, the minimum time between two crises in one country is five years. Thus, by assuming that crises last two years, we bias our noise-to-signal ratio upwards, as only type 2 errors can be issued during the quarters immediately following the end of the second year after an episode.

Consider next the release phase. The approach here is broadly similar but recognizes explicitly that crises are sudden events. The release signal has to be issued as a *coincident*, not leading, indicator of distress. Again, the signal for the release is "on" if the indicator variable breaches a particular threshold. For a variable y_t^l which should increase once crises materialize (for example, credit spreads) the signal is $S(y_t) = 1$ if $y_t^l > k$ & $y_{t-1}^l < k$.[24] To account for the uncertainty in the precise dating of crises, we judge a signal of 1 (0) to be correct if a crisis occurs (does not occur) during a window of three quarters starting with the quarter prior to the date in which the signal is issued. Again, we assess a range of thresholds for each indicator variable and compute type 1 errors, type 2 errors, and the noise-to-signal ratio.

The literature assesses early-warning indicators on the basis of their noise-to-signal ratio (e.g., Kaminsky and Reinhart 1999). However, Demirgüç-Kunt and Detragiache (1999) suggest that this is not ideal from a policy perspective, as policymakers may assign more weight to the risk of missing crises (type 1 error) than to the risk of calling those that do not occur (type 2 error), as the costs of the two

[24] The procedure is symmetric for variables the value of which tends to drop during a crisis (for example, credit growth).

differ.[25] Jordà and Taylor (2011) suggest an alternative approach to capture the trade-off between type 1 and type 2 errors. They construct the correct classification frontier.[26]

As the preferences of policymakers are unobservable, Borio and Drehmann (2009a) suggest minimizing the noise-to-signal ratio subject to at least two-thirds of the crises being correctly predicted. They show that the more concerned a policymaker is about missing crises (type 1 error), the lower are the critical thresholds to be crossed before signaling crises and the noisier the indicators become. At the other end of the spectrum, minimizing the noise-to-signal ratio regardless of the number of crisis predictions generally results in an unacceptably low percentage of crises predicted. On balance, they find that minimizing the noise-to-signal ratio subject to at least two-thirds of the crises being correctly predicted appears to provide a good compromise. We follow this criterion, even though the key messages of the paper would remain unchanged if we had chosen other cut-off levels for the minimum required fraction of correctly predicted crises (e.g., 50 percent or 75 percent). In tables 2–9 (discussed below) we use boldface entries in the columns labeled "Predicted" to denote threshold values that lead to a crisis prediction rate of at least 66 percent. The boldface entries in the columns labeled NS indicate the lowest noise-to-signal ratio for those threshold values that satisfy the condition of a minimum 66 percent prediction rate. Tables 2–5 in the main text only show the best performance of different indicators, i.e., they show the threshold for each indicator which achieves the lowest noise-to-signal ratio whilst capturing at least 66 percent of the crises, or, if this requirement is not fulfilled, simply the threshold with the lowest noise-to-signal ratio. The extended analysis is presented in tables 6–9 in the appendix.

[25]This problem was also highlighted by Bussière and Fratzscher (2008) in the context of logit models. See also Alessi and Detken (2011) for an application and extension.

[26]Similar to a production possibilities frontier, the correct classification frontier plots the relationship between correctly predicting crises and type 2 errors. Given a particular utility function, policymakers could pick the threshold that maximizes their utility. Jordà and Taylor (2011) suggest that, independently of policymakers' preferences, the area below this curve represents a simple non-parametric statistic for the usefulness of an indicator.

Table 2. The Best Performance of Indicator Variables for the Build-Up

	All				Homogenous Data			
Variable	**#Cr**	**TH**	**Pred.**	**NS**	**#Cr**	**TH**	**Pred.**	**NS**
Macroeconomic Variables								
GDP Growth	40	3.5	**90**	**60**	18	3.5	**89**	**39**
GDP Gap	37	3	**76**	**49**	18	3	**78**	**46**
Credit Growth	45	12	**67**	**33**	18	8.5	**67**	**28**
Credit Growth – GDP Growth	39	8	**69**	**23**	18	5	**67**	**37**
Credit-to-GDP Gap	36	10	**67**	**16**	18	12	**67**	**15**
M2 Growth	40	14	**70**	**53**	18	12	**67**	**16**
Property Growth	25	7	**68**	**30**	18	3	**72**	**65**
Property Gap	22	10	**77**	**33**	18	10	**78**	**40**
Equity Growth	28	23	**79**	**34**	18	21	**89**	**32**
Equity Gap	26	10	**69**	**60**	18	20	**67**	**50**
Banking Sector Conditions								
Profits	28	0.70	**71**	**79**	18	0.60	**72**	**92**
Loss	28	0.40	**68**	**77**	18	0.40	**72**	**93**
Market Indicators								
Bank CDS	4	15	**75**	**44**				
LIBOR-OIS	3	10	**67**	**60**				
Credit Spreads	4	130	**75**	**79**				

Notes: #Cr: number of crises that are included in the analysis. TH: threshold; for macroeconomic variables and banking sector conditions in percent; for market indicators in basis points. Pred.: percentage of crises predicted correctly. NS: noise-to-signal ratio; fraction of type 2 errors (a signal is issued and no crisis occurs) divided by one minus the fraction of type 1 errors (no signal is issued and a crisis occurs). For all variables (except profits and market indicators), a signal equal to 1 is issued when the conditioning variable exceeds (is smaller than) the threshold. Otherwise, the signal is equal to 0. A signal of 1 (0) is judged to be correct if a crisis (no crisis) occurs any time within a three-year horizon. Bold figures indicate that more than 66 percent of crises are captured. For a more detailed analysis, see table 6 in the appendix.

5.2 The Build-Up Phase

Table 2 shows the best performance of the different indicator variables for the build-up phase. The results of the full analysis are provided in table 6 in the appendix. For macroeconomic variables and banking sector conditions, the tables show the results for the full and restricted samples, with the latter including only the period for which we have a complete set of variables for all countries.[27] Initially we focus only on the results when all data are used. We first highlight the best-performing indicator and then discuss the remaining ones, ordering them roughly from the worst to the best.

Looking at table 2, we see that the credit-to-GDP gap achieves the lowest noise-to-signal ratio, 16 percent, while still capturing at least 66 percent of the crises in the sample.

The worst performers are indicators of banking sector conditions. As already suggested by the descriptive figures, they cannot signal the build-up of systemic risk well. When at least 66 percent of the crises are captured, the signals are quite unreliable (noise-to-signal ratios well above 70 percent).

Noise-to-signal ratios for market-based indicators are also very high, varying between 44 percent and 80 percent. In contrast to other variables, the results for these indicators highlight a striking pattern, albeit subject to the strong caveat related to the very limited number of crises: the indicators' performance improves dramatically from predicting no crises to capturing all episodes with only small changes in the thresholds. For instance, average CDS spreads across the whole sample are around 50 basis points, peaking well above 100 basis points during the recent crisis. Yet all the crises in the sample (which are only four) would have been correctly signaled by spreads below 20 basis points and none would have been captured by a threshold of 5 basis points. In essence, this means that during normal and boom times, these variables do not seem to fluctuate much. Hence, these series are unable to provide useful signals about the intensity of the build-up of systemic risk.

Output variables, money growth, and the equity gap also perform poorly, albeit slightly better than banking sector conditions and market-based indicators. The lowest noise-to-signal ratios lie between 40 and 60 percent. The results for GDP growth and the output gap argue against linking countercyclical capital buffers to output if the macroprudential objective is defined narrowly in terms of protecting the banking system against the build-up of systemic stress.

---

[27] As additional robustness checks, we analyzed data samples that include only countries that are members of the Basel Committee, or only data from 1980 onwards. We also analyzed a broader range of thresholds than those shown in the tables. These alternatives do not affect the main messages of the analysis. Results are available on request.

The remaining variables (credit growth, the difference between credit and GDP growth, equity price growth, and property price growth and its gap) seem to be the second-best class of indicators after the credit-to-GDP gap. This is not surprising, as these variables capture key aspects of the build-up of systemic risk ahead of many crises, namely an unsustainable credit expansion alongside booming asset prices.

The broad picture is unchanged when the homogenous data set is used. The key difference is that money growth performs nearly as well as the credit-to-GDP gap in this sample.[28] But in contrast to the case of the credit-to-GDP gap, this result is not stable. When we consider the entire data set (and, hence, all crisis episodes) and select thresholds that allow the indicators to capture more than two-thirds of crises, the noise-to-signal ratio increases unacceptably to above 50 percent.

The results provide valuable information about the performance of different variables. However, we need to keep two caveats in mind.

First, we have assessed the signaling properties of *domestic* indicators of the financial cycle. But systemic problems may occur because of banks' foreign exposures. The cases of German and Swiss banks in the recent crisis are obvious examples. Borio and Drehmann (2009a) show how this signaling problem can be partially addressed by incorporating foreign claims in the assessment of banks' vulnerabilities. We explore this in more detail in section 6.2.

Second, a statistical type 2 error is not necessarily a type 2 error from a policy perspective. Often the conditioning variable starts indicating the build-up of vulnerabilities earlier than three years before a crisis. In the statistical analysis, such a signal is counted as false even though it provides the right information, but simply "too early." There are also instances of severe banking strains without a crisis being formally recorded (possibly because of mitigating policy action to diffuse pressure on the banking system). Therefore, an indicator may issue false signals in the statistical sense, even though additional capital buffers would have been highly valuable to cushion the impact of the stress on the banking system.

[28] The performance of property prices is also sensitive to the sample. This partly reflects the requirement that 66 percent of the crises should be predicted. A one-to-one comparison for each threshold indicates a slightly weaker predictive power for each threshold with a similar level of type 2 errors, resulting in somewhat higher noise-to-signal ratios.

Even with this broader view of false signals, it is clear that no variable provides perfect signals. This means that, in practice, pure rule-based schemes may not be desirable. Some form of discretion may prove inevitable.

5.3 The Release Phase

Although the credit-to-GDP gap is the best-performing indicator for the build-up phase, figure 2 indicates that it declines only slowly once crises materialize. This is also borne out by the statistical tests shown in table 3 (the results of the full analysis are provided in table 7 in the appendix). As before, bold values for "Predicted" highlight thresholds for which a release signal is issued correctly for at least 66 percent of the crises. The bold noise-to-signal ratio indicates the lowest noise-to-signal ratio for all threshold values that satisfy this condition.

None of the macro variables and of the indicators of banking sector conditions satisfy the required degree of predictive power to make them robust anchor variables for the release phase; i.e., none of these variables signal more than 66 percent of the crises. The best indicator is a drop of credit growth below 8 percent. This happens at the onset of more than 40 percent of crises, and such a signal provides very few false alarms (the noise-to-signal ratio is around 10 percent).

Market-based indicators do signal the onset of crises but with considerable noise (in terms of false signals). Take credit spreads, for which most data are available. They breach the 200-basis-point barrier in 75 percent of all crisis episodes. However, the noise-to-signal ratio is close to unity, rendering the overall signal unreliable. This partly reflects the high correlation of spreads across countries. The 200-basis-point threshold was breached in Canada and Australia around 2007:Q3, even though neither country experienced a crisis. Credit spreads also rose quickly during the dot-com bust, a period not associated with banking crises. That said, all these results are derived from a sample that is too small to support robust conclusions.

Overall, the results indicate that policymakers may need to rely much more on discretion for the release phase than for the build-up phase. No single variable provides reliable and robust signals for this stage.

Table 3. The Best Performance of Indicator Variables for the Release

	All				Homogenous Data			
Variable	**#Cr**	**TH**	**Pred.**	**NS**	**#Cr**	**TH**	**Pred.**	**NS**
Macroeconomic Variables								
GDP Growth	40	5	25	25	18	5	6	87
GDP Gap	37	3	27	12	18	3	28	11
Credit Growth	45	8	43	12	18	8	44	9
Credit Growth – GDP Growth	39	6	26	22	18	6	28	18
Credit-to-GDP Gap	36	10	14	11	18	10	22	8
M2 Growth	40	12	20	25	18	12	17	20
Property Growth	25	2	12	42	18	2	11	35
Property Gap	22	4	32	6	18	4	39	3
Equity Growth	28	23	21	35	18	23	22	30
Equity Gap	26	15	19	21	18	15	22	20
Banking Sector Conditions								
Profits	28	0.60	31	24	18	0.60	30	30
Loss	28	0.70	32	229	18	0.70	33	171
Market Indicators								
Bank CDS	4	10	50	156				
LIBOR-OIS	3	50	**67**	**90**				
Credit Spreads	4	210	**75**	**105**				

Notes: #Cr: number of crises that are included in the analysis. TH: threshold; for macroeconomic variables and banking sector conditions in percent; for market indicators in basis points. Pred.: percentage of crises predicted correctly. NS: noise-to-signal ratio; fraction of type 2 errors (a signal is issued and no crisis occurs) divided by one minus the fraction of type 1 errors (no signal is issued and a crisis occurs). For all variables (except profits and market indicators), a signal equal to 1 is issued when the conditioning variable exceeds (is smaller than) the threshold. Otherwise, the signal is equal to 0. A signal of 1 (0) is judged to be correct if a crisis (no crisis) occurs any time within a three-year horizon. Bold figures indicate that more than 66 percent of crises are captured. For a more detailed analysis, see table 7 in the appendix.

6. Robustness

In this section we evaluate the robustness of our results. First, we assess the signaling properties of the credit-to-GDP gap calculated by reference to alternative specifications of the trend. We consider different specifications for the one-sided HP filter and we discuss a more conventional two-sided filter. Second, we explore different signaling horizons. Third, we explore how our analysis linking domestic indicator variables with banking crises gels with a financial system that is global and exposes banks to shocks from abroad.

6.1 The Signaling Properties of Different Credit-to-GDP Gaps

In this section we assess the impact of two unconventional choices we made in the calculation of the credit-to-GDP trend. The first relates to the choice of a one-sided trend (i.e., a backward-looking Hodrick-Prescott filter) and the second to that of the smoothing parameter λ.

Edge and Meisenzahl (2011) perform an analysis for the credit-to-GDP ratio using real-time and revised data for the United States and conclude that revisions can lead to substantial changes to the estimated gap.[29] They evaluate the impact of revisions to the credit-to-GDP gap from two sources: data revisions and the unfolding of history. As regards the former, they compare the impact of revisions to the credit and GDP series on the original estimate of trend and find that they contribute only to a small extent to the revised gap estimates. As regards the latter, they compare the one-sided with a two-sided filter, which encompasses information *about the future* relative to the point in time when decisions are taken. They find that the one-sided trend differs substantially from the two-sided one. Based on this, they conclude that the credit-to-GDP gap is an unreliable guide for countercyclical capital buffers.

The message of Edge and Meisenzahl (2011) is, in our view, misleading. They focus their attention on the difference between the gap calculated using the one-sided filter ending in a given quarter to that calculated using information from subsequent quarters. The mismeasurement they identify is obviously *impossible* for the policymaker to correct in real time, *since the data needed cannot be available*. Moreover, and most importantly for the purpose of this paper, they do not assess the indicator performance of the credit-to-GDP gap calculated on the basis of future information. Only if the indicator performance is seriously hampered by the calculation of the trend could their conclusion be a reason for concern. Even then, from an applied policy perspective, a trend calculation that requires future information is problematic.

[29] A similar analysis was conducted for the U.S. GDP gap by Orphanides and van Norden (2002).

Table 4. The Best Performance of Different Credit-to-GDP Gaps

	All				Homogenous Data			
Horizon	**#Cr**	**TH**	**Pred.**	**NS**	**#Cr**	**TH**	**Pred.**	**NS**
Real Time, HP $\lambda = 400{,}000$ (Standard)	36	10	**67**	**16**	18	12	**67**	**15**
Two-Sided Filter (Past and Future Known), HP $\lambda = 400{,}000$	36	6	**72**	**17**	18	8	**67**	**10**
Real Time, HP $\lambda = 1{,}600$	36	2.5	**69**	**30**	18	2	**67**	**46**
Real Time, Linear Trend	36	8	**67**	**22**	18	10	**67**	**23**

Notes: #Cr: number of crises that are included in the analysis. TH: threshold; in percent. Pred.: percentage of crises predicted correctly. NS: noise-to-signal ratio; fraction of type 2 errors (a signal is issued and no crisis occurs) divided by one minus the fraction of type 1 errors (no signal is issued and a crisis occurs). For all variables, a signal equal to 1 is issued when the conditioning variable exceeds the threshold. Otherwise, the signal is equal to 0. A signal of 1 (0) is judged to be correct if a crisis (no crisis) occurs any time within a three-year horizon. Bold figures indicate that more than 66 percent of crises are captured. Real-time trends use information available up to each point in time in which the signal is issued. The two-sided filter uses all available information in the data set. For a more detailed analysis, see table 8 in the appendix.

The first and second rows in table 4 (see table 8 in the appendix for a more detailed analysis) compare the performance of the gap based on the one-sided (labeled "standard") and two-sided Hodrick-Prescott (HP) filters (the results of the full analysis are provided in table 8 in the appendix). Both trends are estimated using a smoothing factor λ of 400,000. The table shows that knowing the future does not actually help in this case. For the homogenous data set, the statistical performance is marginally better. But it is actually worse if all data are considered.

The second robustness issue refers to the choice of λ. First, in line with standard business-cycle analysis, we construct credit-to-GDP gaps using a one-sided HP filter with a smoothing factor λ of 1,600.[30] As discussed above, this implies that credit cycles would have the

[30] As a robustness check, we also analyzed output and asset price gaps using a smoothing factor λ of 1,600. This does not improve the performance of these variables in comparison to the results shown in table 6. Drehmann et al. (2010) also assess credit-to-GDP gaps based on $\lambda = 25{,}000$ and $\lambda = 125{,}000$, which

same length as business cycles. Second, we use a simple linear time trend, based on fifteen years of rolling regressions.[31]

Table 4 (and the full results in table 8 in the appendix) shows that a linear trend performs well, even though it is slightly noisier than the credit-to-GDP gap based on an HP filter with $\lambda = 400{,}000$. The table illustrates that a gap calculated using $\lambda = 1{,}600$ (i.e., assuming that credit cycles and business cycles are of the same length) performs poorly. We take this result as another indication that the financial and business cycles are not the same.

6.2 The Forecast Horizon

As discussed in the methodology section 5.1, in the benchmark specification we adopt a flexible forecast horizon, so that a crisis signal is judged to be correct (false) if a crisis (no crisis) occurs *any time* within a three-year interval. As already argued, in our view this is the right approach, as indicators only highlight the build-up of vulnerabilities rather than provide clear-cut signals of the precise future timing of crises. Even so, here we explore different forecast horizons. Table 5 shows the key results (the results of the full analysis are provided in table 9 in the appendix) when a crisis signal is judged to be correct if a crisis occurs in the next, the second, *or* the third year ahead. In addition, we analyze two-year horizons for years 1 and 2, and 2 and 3, respectively. For brevity we only show the results for the credit-to-GDP gap, which remains the best single indicator for all different horizons considered.[32]

When the forecast horizon is limited to a single year, the performance deteriorates relative to our standard approach, but not by too much. This is not surprising, as figure 2 reveals that the credit-to-GDP gap is highly persistent. Take a threshold of 4 percentage points (see table 9 in the appendix). In the standard approach with the homogenous data set, 89 percent of the crises are captured and the

---

assumes that credit cycles are two or three years longer than business cycles. The statistical performance of the credit gap with $\lambda = 400{,}000$ remains best.

[31] We also used five- or ten-year windows to construct the trend. Predictably, decreasing the number of years used to construct the linear trend worsens the performance of the gap as indicator. In particular, during periods of sustained credit growth, the trend catches up too quickly, so that gaps start declining more markedly ahead of crises.

[32] Results for other variables are available on request.

Table 5. The Best Performance of Different Signaling Horizons

	All				Homogenous Data			
Horizon	**#Cr**	**TH**	**Pred.**	**NS**	**#Cr**	**TH**	**Pred.**	**NS**
Year 1 & 2 & 3 (Standard)	36	10	**67**	**16**	18	12	**67**	**15**
Year 1	36	6	**67**	**36**	18	11	**67**	**24**
Year 2	36	6	**67**	**34**	18	9	**67**	**28**
Year 3	36	5	**69**	**36**	18	5	**78**	**40**
Year 1 & 2	36	8	**67**	**25**	18	11	**67**	**21**
Year 2 & 3	36	7	**67**	**27**	18	10	**72**	**19**

Notes: #Cr: number of crises that are included in the analysis. TH: threshold; in percent. Pred.: percentage of crises predicted correctly. NS: noise-to-signal ratio; fraction of type 2 errors (a signal is issued and no crisis occurs) divided by one minus the fraction of type 1 errors (no signal is issued and a crisis occurs). A signal equal to 1 is issued when the credit-to-GDP gap exceeds the threshold. Otherwise, the signal is equal to 0. A signal of 1 at time t is judged to be correct (false) if a crisis (no crisis) occurs over the following signaling horizons: (i) "year 1 to 3": quarters $t+1$ to $t+12$, (ii) "year 1": quarters $t+1$ to $t+4$, (iii) "year 2": quarters $t+5$ to $t+8$, (iv) "year 3": quarters $t+9$ to $t+12$, (v) "year 1 and 2": quarters $t+1$ to $t+8$, and (vi) "year 2 and 3": quarters $t+5$ to $t+12$. Bold figures indicate that more than 66 percent of crises are captured. For a more detailed analysis, see table 9 in the appendix.

type 2 error is 36 percent. When only individual years are considered, the predictive power drops to 83 percent and the noise increases (the worst case is for year 1, where type 2 errors are equal to 41 percent). These effects are somewhat more pronounced when all data are considered.

Interestingly, in comparison to the standard approach, the predictive power is only marginally reduced when a two-year horizon is used. And in this case it is virtually the same when the signaling horizon covers years 1 and 2 or 2 and 3. From an operational perspective, the strong performance of the credit-to-GDP gap in years 2 and 3 before crises is important, as capital planning by the banks requires that they know their regulatory capital requirement at least one year before they become effective (Basel Committee on Banking Supervision 2010). Hence, if countercyclical capital requirements were raised following a signal, capital would be available only with a one-year lag.

6.3 The International Dimension

Finance is increasingly international. The outstanding stock of banks' foreign claims tripled between 2000 and 2007 from $11 trillion to over $30 trillion. The ten largest banks operate on average in around eighty countries in the world. Cross-border borrowing and lending has implications for the specification of the anchor variable. For one, domestic borrowers' improved access to international sources of credit means that domestic credit figures may understate the extent of their leverage and vulnerability to shocks. In addition, banks located in a given country may be exposed to risks in other countries—risks that are not captured by a domestically focused indicator. Crises can occur because of losses on foreign exposures. In this section we discuss how the credit-to-GDP ratio can be adjusted to deal with these issues.

The first complication relates to the coverage of the credit variable proxying the financial cycle *in a given country.* It covers lending to households and business residents in a given jurisdiction. The reliability of this proxy should be, as a first approximation, independent of the source of credit (domestic, international, bank originated, or market based). Hence, the credit aggregate should be as broad as possible. In our analysis, however, the credit series generally include only credit granted by banks *located* in the given country; i.e., they exclude direct cross-border lending to domestic households and businesses from non-residents. This is typically what the credit and banking statistics cover. We have conducted a preliminary analysis of the effect of including a more comprehensive measure of credit to domestic borrowers, drawing on the BIS international banking statistics.[33] The preliminary results are consistent with those reported in this paper. Large credit-to-GDP gaps provide a reliable signal ahead of systemic banking crises.

The second complication relates to the possibility of exposures of the banking system in a given country to financial cycles in the rest

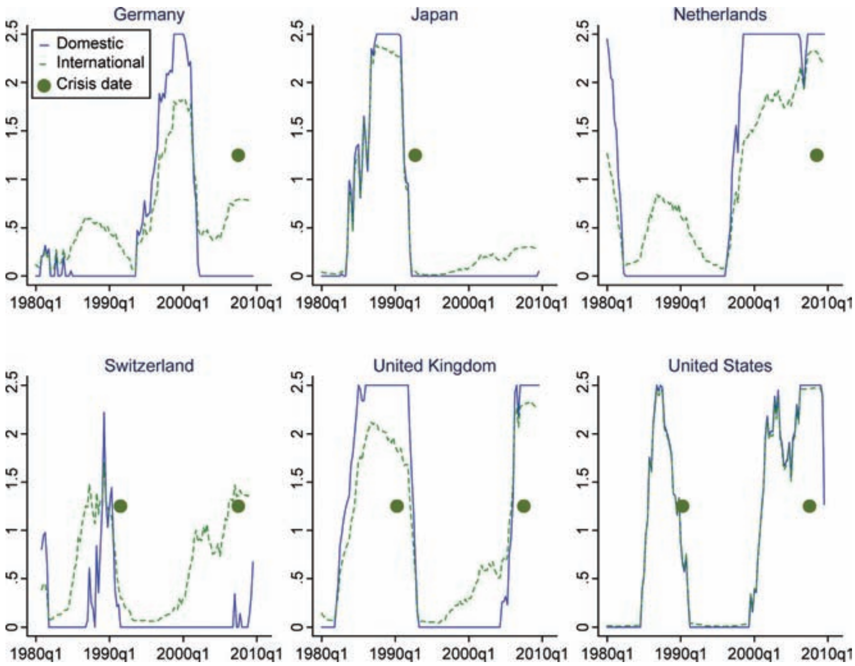
[33] A number of data issues complicate this analysis. For instance, data availability on non-bank sources of finance is rather poor in some countries, and differences in the treatment of exchange rate changes in the reported series can introduce significant noise.

of the world. This means that, in principle, the signaling property of the credit-to-GDP indicator should be evaluated on the basis of the composition of *global* exposures of domestic banks. While these are known to supervisors, in our analysis we can only use rough aggregate proxies. To do so, we follow a three-step procedure similar to that in Borio and Drehmann (2009b). We first need to convert the indicator into something that can be combined across economies. We do this by constructing a (hypothetical) capital surcharge that varies with the value of the credit-to-GDP gap. The mapping we selected is one in which the surcharge increases linearly between zero and 2.5 percent as the value of the indicator increases between a minimum level (equal to 2 percentage points) and a maximum level (equal to 10 percentage points). So the surcharge is zero if the credit-to-GDP gap is less than 2 percentage points, and it is equal to 2.5 percent if the gap is greater than or equal to 10 percentage points. This calibration is in line with the proposals by the Basel Committee on Banking Supervision (2010). In the second step, we calculate for each country the share of total assets of banks headquartered in that country that is accounted for by claims vis-à-vis other countries. This is done by drawing on the BIS international banking statistics.[34] In the final step, we use these shares to construct a weighted average of the capital charges that would apply to exposures in each country in which the banks have a claim.

Figure 5 compares for six countries (Germany, Japan, the Netherlands, Switzerland, the United Kingdom, and the United States) the capital surcharge calculated based solely on domestic exposures with that based on the global exposures of the representative bank. The dots denote banking crises. The results are encouraging. Capital buffers build up to their maximum ahead of major systemic crises. Moreover, the indicator that includes proxies for banks' international exposures captures better the problems faced by the German and Swiss banking sectors, which incurred losses largely on lending to non-residents, notably to U.S. borrowers.

[34] The statistics refer to aggregate claims and not individual bank claims. They are reported on a consolidated basis. For a description of structure and coverage of the statistics, see McGuire and Wooldridge (2005).

Figure 5. International Exposures and Countercyclical Capital Charges



Notes: The domestic buffer reflects the credit-to-GDP gap of the specific country and would reflect the total charge for banks located there if they did not have any international exposures. The international buffer is the buffer of a hypothetical bank whose portfolio of domestic and cross-border credit corresponds to that of the banking system in the given jurisdiction. Country exposure weights are based on the BIS international banking statistics and correspond to claims as of 2006:Q4.

7. Conclusion

The objective of prudential countercyclical bank capital standards is to encourage banks to increase their defenses against systemwide vulnerabilities by building up buffers in good times so that they can draw them down in bad times. Our analysis examined the suitability of different variables to act as anchors for the build-up and release phases of the buffers.

The analysis shows that the best variables to signal the pace and size of the *build-up* of the buffers differ from those that provide the best signals for their *release*. Credit, measured by the deviation of the credit-to-GDP ratio from its trend, emerges as the best variable for the build-up phase, as it has the strongest leading indicator properties for financial system distress. A side benefit of using this variable as the anchor is that it could help to restrain the credit boom and hence risk taking to some extent.

How to guide the pace and intensity of the release of the buffer is less clear. In general, a prompt and sizable release is desirable, as a gradual release could reduce the buffer's effectiveness. A combination of a measure of aggregate losses with indicators of tightening of credit conditions would provide, conceptually, the proper signal for the beginning of the period of systemwide stress, thus triggering the release. Among the single indicators we evaluated, credit spreads are the most promising, albeit over the shorter period over which we were able to assess them. But the performance of these variables for the release is not as good as that of the credit-to-GDP gap for the build-up.

Our analysis also makes clear that any operational framework would need to incorporate an element of judgment, especially in the release phase. As in other fields of economic policy, rules provide invaluable discipline but may not work well in all circumstances. Given the relatively early stage in the economic analysis of the interactions between the real and financial sectors of the economy, it would be premature to claim that any rule can be sufficiently robust across countries and time. Moreover, the political economy of the design and application of macroprudential instruments, such as the countercyclical capital buffer, is a field in which much more analysis is needed.

A final word of caution is in order. Are our empirical results subject to the usual Lucas or Goodhart critiques? In other words, *if the scheme proved successful*, would the leading indicator properties of the credit-to-GDP variable disappear? The answer is "yes," by definition, if the criterion of success was avoiding major distress among banks. As credit exceeds the critical threshold, banks would build up buffers to withstand the bust. Moreover, if, in addition, the scheme acted as a brake on risk taking during the boom, the bust would be less likely in the first place. The answer is less clear if the criterion was the more ambitious one of avoiding disruptive financial busts: busts could occur even if *banks* remained reasonably resilient. In either situation, however, the loss of predictive content per se would be *no reason* to abandon the scheme.

Appendix

Table 6. The Performance of Macroeconomic Variables for the Build-Up

Threshold	All Data				Homogenous Data			
	T 1	**T 2**	**Pred.**	**NS**	**T 1**	**T 2**	**Pred.**	**NS**
GDP Growth (40/18)								
3	8	61	**93**	66	11	42	**89**	48
3.5	10	54	**90**	**60**	11	35	**89**	**39**
4	23	48	**78**	62	33	28	**67**	42
4.5	33	42	**68**	62	50	21	50	41
5	35	36	65	56	56	15	44	35
GDP Gap (37/18)								
1	16	53	**84**	63	11	54	**89**	61
2	16	44	**84**	53	11	44	**89**	50
3	24	37	**76**	**49**	22	36	**78**	**46**
4	35	30	65	47	44	29	56	52
Credit Growth (45/18)								
8	16	40	**84**	47	22	23	**78**	29
8.5	20	37	**80**	46	33	19	**67**	**28**
9	22	34	**78**	44	39	16	61	26
11	29	26	**71**	36	44	8	56	15
12	33	22	**67**	**33**	50	6	50	12
13	38	19	62	30	56	5	44	12
Credit Growth – GDP Growth (39/18)								
4	10	38	**90**	42	11	32	**89**	36
5	21	30	**79**	38	33	24	**67**	**37**
6	23	24	**77**	32	39	17	61	28
8	31	16	**69**	**23**	44	9	56	16
9	38	13	62	21	44	6	56	11
Credit-to-GDP Gap (36/18)								
2	11	38	**89**	43	6	45	**94**	48
4	17	29	**83**	35	11	36	**89**	41
6	25	21	**75**	28	17	27	**83**	33
8	31	16	**69**	23	22	20	**78**	26
10	33	11	**67**	**16**	22	14	**78**	18
12	47	8	53	14	33	10	**67**	**15**
14	58	5	42	13	50	7	50	15
M2 Growth (40/18)								
8	10	64	**90**	71	17	30	**83**	36
10	18	53	**83**	65	22	17	**78**	22
12	23	44	**78**	57	33	11	**67**	**16**
14	30	37	**70**	**53**	44	8	56	14
16	40	31	60	52	67	6	33	18

(continued)

Table 6. (Continued)

Threshold	All Data				Homogenous Data			
	T 1	**T 2**	**Pred.**	**NS**	**T 1**	**T 2**	**Pred.**	**NS**
Property Growth (25/18)								
2	16	51	**84**	61	22	53	**78**	68
3	20	45	**80**	56	28	47	**72**	**65**
4	28	39	**72**	54	39	41	61	66
6	28	26	**72**	37	39	29	61	47
7	32	20	**68**	**30**	39	21	61	34
8	44	17	56	30	56	17	44	38
Property Gap (22/18)								
6	14	38	**86**	43	17	43	**83**	52
8	18	31	**82**	38	22	38	**78**	49
10	23	26	**77**	**33**	22	31	**78**	**40**
12	36	22	64	35	39	27	61	43
Equity Growth (28/18)								
19	14	32	**86**	37	11	31	**89**	35
21	14	30	**86**	34	11	28	**89**	**32**
23	21	27	**79**	**34**	22	26	**78**	33
25	39	25	61	42	44	24	56	44
Equity Gap (26/18)								
5	19	49	**81**	61	11	54	**89**	61
10	31	42	**69**	**60**	22	46	**78**	59
15	35	35	65	53	28	40	**72**	55
20	42	29	58	50	33	33	**67**	**50**
25	58	24	42	58	56	28	44	64
Profits (28/18)								
0.40	4	82	**96**	85	6	84	**94**	89
0.50	11	76	**89**	85	17	78	**83**	94
0.60	21	67	**79**	86	28	66	**72**	**92**
0.70	29	57	**71**	**79**	39	58	61	95
0.80	46	47	54	87	56	50	44	113
0.90	57	38	43	89	56	41	44	93
Loss (28/18)								
0.70	4	80	**96**	83	6	89	**94**	94
0.60	4	73	**96**	76	6	85	**94**	90
0.50	11	63	**89**	71	11	79	**89**	88
0.40	32	52	**68**	**77**	28	67	**72**	**93**
0.30	50	39	50	78	39	54	61	89
0.20	61	24	39	61	50	38	50	76

(continued)

Table 6. (Continued)

Threshold	All Data				Homogenous Data			
	T 1	T 2	Pred.	NS	T 1	T 2	Pred.	NS
Bank CDS (4/0)								
20	0	49	**100**	49				
15	25	33	**75**	**44**				
10	50	8	50	16				
5	100	0	0	0				
LIBOR-OIS (3/0)								
14	33	73	**67**	109				
12	33	55	**67**	82				
10	33	40	**67**	**60**				
8	67	37	33	111				
Credit Spreads (4/0)								
140	25	62	**75**	82				
130	25	59	**75**	**79**				
120	50	54	50	109				
110	50	49	50	97				

Notes: Threshold: for macroeconomic variables and banking sector conditions in percent; for market indicators in basis points. T1: type 1 error, no signal is issued and a crisis occurs. T2: type 2 error, a signal is issued and no crisis occurs. Pred.: percentage of crises predicted correctly. Bold figures in this column indicate that more than 66 percent of crises are captured. NS: noise-to-signal ratio; fraction of type 2 errors divided by one minus the fraction of type 1 errors. Bold figures in this column indicate the lowest noise-to-signal ratio given that more than 66 percent of crises are captured. For all variables (except profits and market indicators), a signal equal to 1 is issued when the conditioning variable exceeds (is smaller than) the threshold. Otherwise, the signal is equal to 0. A signal of 1 (0) is judged to be correct if a crisis (no crisis) occurs any time within a three-year horizon. The numbers in parentheses next to the name of the variable correspond to the number of crises that are included in the analysis using the sample that includes the available data for that variable (left-hand-side number) and the homogenous sample (right-hand-side number), respectively.

Table 7. The Performance of Indicator Variables for the Release

Threshold	All Data				Homogenous Data			
	T 1	**T 2**	**Pred.**	**NS**	**T 1**	**T 2**	**Pred.**	**NS**
GDP Growth (40/18)								
3	78	7	23	29	94	8	6	140
3.5	80	7	20	34	100	7	0	—
4	73	7	28	26	94	7	6	130
4.5	78	7	23	30	94	6	6	114
5	75	6	25	25	94	5	6	87
GDP Gap (37/18)								
1	78	4	22	17	67	3	33	9
2	73	4	27	13	67	3	33	8
3	73	3	27	12	72	3	28	11
4	76	3	24	11	89	3	11	24
5	89	2	11	22	94	3	6	47
Credit Growth (45/18)								
6	67	6	33	17	67	5	33	16
8	57	5	43	12	56	4	44	9
10	70	5	30	15	72	3	28	10
12	78	4	22	20	89	2	11	16
14	83	4	17	21	94	1	6	23
Credit Growth – GDP Growth (39/18)								
0	82	6	18	34	83	5	17	32
2	85	7	15	47	78	7	22	30
4	74	7	26	26	72	5	28	20
6	74	6	26	22	72	5	28	18
8	79	4	21	21	83	3	17	20
Credit-to-GDP Gap (36/18)								
2	86	3	14	24	78	3	22	12
4	94	3	6	47	89	2	11	22
6	94	2	6	38	89	2	11	18
8	94	2	6	31	89	2	11	18
10	86	2	14	11	78	2	22	8
M2 Growth (40/18)								
8	75	6	25	23	72	7	28	27
10	78	6	23	25	78	5	22	23
12	80	5	20	25	83	3	17	20
14	88	5	13	37	94	3	6	51

(continued)

Table 7. (Continued)

Threshold	All Data				Homogenous Data			
	T 1	**T 2**	**Pred.**	**NS**	**T 1**	**T 2**	**Pred.**	**NS**
Property Growth (25/18)								
2	88	5	12	42	89	4	11	35
3	88	5	12	44	89	4	11	36
4	92	5	8	65	94	4	6	69
5	96	5	4	124	100	4	0	—
6	92	5	8	61	100	4	0	—
Property Gap (22/18)								
2	68	2	32	6	67	2	33	5
4	68	2	32	6	61	1	39	3
6	73	2	27	8	67	2	33	6
8	82	2	18	11	78	2	22	9
10	82	2	18	10	78	2	22	9
Equity Growth (28/18)								
15	82	9	18	50	83	9	17	53
17	82	9	18	48	83	8	17	51
19	82	9	18	49	83	8	17	51
21	79	8	21	38	78	8	22	35
23	79	7	21	35	78	7	22	30
Equity Gap (26/18)								
0	77	5	23	22	78	5	22	21
5	81	5	19	26	83	5	17	30
10	85	5	15	31	83	5	17	31
15	81	4	19	21	78	5	22	20
20	85	4	15	25	83	4	17	25
Profits (28/18)								
0.40	75	6	25	25	78	8	22	35
0.50	65	7	35	20	74	9	26	34
0.60	69	7	31	24	70	9	30	30
0.70	71	8	29	29	74	8	26	30
0.80	76	8	24	32	78	8	22	35
0.90	83	8	17	46	81	9	19	50
Loss (28/18)								
0.70	68	74	32	229	67	57	33	171
0.60	68	78	32	243	67	67	33	200
0.50	68	81	32	252	63	69	37	187
0.40	73	82	27	301	67	75	33	225
0.30	80	83	20	411	74	75	26	289

(continued)

Table 7. (Continued)

Threshold	All Data				Homogenous Data			
	T 1	**T 2**	**Pred.**	**NS**	**T 1**	**T 2**	**Pred.**	**NS**
Bank CDS (4/0)								
5	100	0	0	0				
10	50	78	50	156				
15	100	100	0	0				
20	75	86	25	343				
LIBOR-OIS (3/0)								
10	33	75	**67**	113				
30	33	60	**67**	90				
50	33	60	**67**	90				
70	33	71	**67**	107				
75	100	100	0	0				
Credit Spreads (4/0)								
160	50	88	50	176				
180	50	89	50	178				
190	25	84	**75**	112				
200	25	83	**75**	111				
210	25	79	**75**	105				
220	50	87	50	173				

Notes: Threshold: for macroeconomic variables and banking sector conditions in percent; for market indicators in basis points. T1: type 1 error, no signal is issued and a crisis occurs. T2: type 2 error, a signal is issued and no crisis occurs. Pred.: percentage of crises predicted correctly. Bold figures in this column indicate that more than 66 percent of crises are captured. NS: noise-to-signal ratio; fraction of type 2 errors divided by one minus the fraction of type 1 errors. Bold figures in this column indicate the lowest noise-to-signal ratio given that more than 66 percent of crises are captured. For all variables (except profits and market indicators), a signal equal to 1 is issued when the conditioning variable exceeds (is smaller than) the threshold. Otherwise, the signal is equal to 0. A signal of 1 (0) is judged to be correct if a crisis (no crisis) occurs any time within a three-year horizon. The numbers in parentheses next to the name of the variable correspond to the number of crises that are included in the analysis using the sample that includes the available data for that variable (left-hand-side number) and the homogenous sample (right-hand-side number), respectively.

Table 8. The Performance of Different Credit-to-GDP Gaps

Threshold	All Data (36)				Homogenous Data (18)			
	T 1	T 2	Pred.	NS	T 1	T 2	Pred.	NS
Real Time, HP $\lambda = 400{,}000$ (Standard)								
2	11	38	**89**	43	6	45	**94**	48
4	17	29	**83**	35	11	36	**89**	41
6	25	21	**75**	28	17	27	**83**	33
8	31	16	**69**	23	22	20	**78**	26
10	33	11	**67**	**16**	22	14	**78**	18
12	47	8	53	14	33	10	**67**	**15**
Two-Sided Filter (Past and Future Known), HP $\lambda = 400{,}000$								
2	17	27	**83**	32	11	24	**89**	27
4	25	18	**75**	24	17	17	**83**	20
6	28	12	**72**	**17**	22	11	**78**	14
8	39	8	61	14	33	7	**67**	**10**
10	50	6	50	11	44	4	56	8
12	61	4	39	9	56	3	44	7
Real Time, HP $\lambda = 1{,}600$								
1	14	41	**86**	47	17	46	**83**	55
1.5	19	33	**81**	41	28	37	**72**	52
2	22	26	**78**	34	33	31	**67**	**46**
2.5	31	21	**69**	**30**	39	24	61	40
3	39	16	61	26	39	19	61	31
3.5	39	12	61	20	39	14	61	23
4	47	9	53	18	50	11	50	22
Real Time, Linear Trend								
4	17	29	**83**	35	6	37	**94**	40
6	22	22	**78**	28	11	29	**89**	32
8	33	15	**67**	**22**	22	21	**78**	28
9	36	12	64	19	22	19	**78**	24
10	44	10	56	18	33	15	**67**	**23**
12	58	7	42	16	44	11	56	19

Notes: Threshold: in percent. T1: type 1 error, no signal is issued and a crisis occurs. T2: type 2 error, a signal is issued and no crisis occurs. Pred.: percentage of crises predicted correctly. Bold figures in this column indicate that more than 66 percent of crises are captured. NS: noise-to-signal ratio; fraction of type 2 errors divided by one minus the fraction of type 1 errors. Bold figures in this column indicate the lowest noise-to-signal ratio given that more than 66 percent of crises are captured. For all variables, a signal equal to 1 is issued when the conditioning variable exceeds the threshold. Otherwise, the signal is equal to 0. A signal of 1 (0) is judged to be correct if a crisis (no crisis) occurs any time within a three-year horizon. The numbers in parentheses next to "All Data" and "Homogenous Data" correspond to the number of crises that are included in the analysis using the respective sample. Real-time trends use information available up to each point in time in which the signal is issued. The two-sided filter uses all available information in the data set.

Table 9. The Performance of Different Signaling Horizons

Threshold	All Data (36)				Homogenous Data (18)			
	T 1	T 2	Pred.	NS	T 1	T 2	Pred.	NS
Year 1 & 2 & 3 (Standard)								
4	17	29	**83**	35	11	36	**89**	41
5	19	25	**81**	31	17	31	**83**	37
6	25	21	**75**	28	17	27	**83**	33
7	31	18	**69**	26	22	23	**78**	30
8	31	16	**69**	23	22	20	**78**	26
9	31	13	**69**	19	22	17	**78**	21
10	33	11	**67**	**16**	22	14	**78**	18
11	39	9	61	15	28	12	**72**	17
12	47	8	53	14	33	10	**67**	**15**
Year 1								
4	25	32	**75**	43	17	41	**83**	50
5	31	28	**69**	40	22	36	**78**	47
6	33	24	**67**	**36**	28	32	**72**	45
7	39	21	61	34	33	28	**67**	43
8	39	18	61	30	33	25	**67**	38
9	39	15	61	25	33	21	**67**	32
10	39	13	61	21	33	18	**67**	28
11	42	11	58	19	33	16	**67**	**24**
12	50	9	50	19	39	14	61	23
Year 2								
4	22	31	**78**	39	17	39	**83**	47
5	28	26	**72**	36	22	34	**78**	43
6	33	23	**67**	**34**	22	30	**78**	38
7	36	19	64	31	22	26	**78**	33
8	39	17	61	28	28	23	**72**	31
9	44	14	56	25	33	19	**67**	**28**
10	50	12	50	24	39	16	61	27
11	58	10	42	24	50	14	50	28
12	61	8	39	22	50	12	50	24
Year 3								
4	28	29	**72**	40	17	36	**83**	44
5	31	25	**69**	**36**	22	31	**78**	**40**
6	42	21	58	37	33	27	**67**	41
7	44	18	56	33	39	23	61	38
8	44	16	56	28	39	20	61	33
9	47	13	53	24	44	17	56	30
10	53	11	47	23	44	14	56	25
11	61	9	39	23	50	12	50	24
12	64	8	36	21	50	10	50	20

(continued)

Table 9. (Continued)

Threshold	All Data (36)				Homogenous Data (18)			
	T 1	**T 2**	**Pred.**	**NS**	**T 1**	**T 2**	**Pred.**	**NS**
Year 1 & 2								
4	19	31	**81**	38	11	39	**89**	44
5	28	26	**72**	36	22	34	**78**	43
6	28	23	**72**	31	22	30	**78**	38
7	33	19	**67**	29	22	26	**78**	33
8	33	17	**67**	**25**	22	23	**78**	29
9	36	14	64	22	28	19	**72**	26
10	39	12	61	19	33	16	**67**	24
11	42	10	58	17	33	14	**67**	**21**
12	50	8	50	17	39	12	61	19
Year 2 & 3								
4	17	29	**83**	35	11	36	**89**	41
5	19	25	**81**	31	17	31	**83**	37
6	31	21	**69**	31	17	27	**83**	33
7	33	18	**67**	**27**	22	23	**78**	30
8	36	16	64	24	28	20	**72**	28
9	36	13	64	20	28	17	**72**	23
10	42	11	58	19	28	14	**72**	**19**
11	56	9	44	20	44	12	56	22
12	58	8	42	18	44	10	56	18

Notes: Threshold: in percent. T1: type 1 error, no signal is issued and a crisis occurs. T2: type 2 error, a signal is issued and no crisis occurs. Pred.: percentage of crises predicted correctly. Bold figures in this column indicate that more than 66 percent of crises are captured. NS: noise-to-signal ratio; fraction of type 2 errors divided by one minus the fraction of type 1 errors. Bold figures in this column indicate the lowest noise-to-signal ratio given that more than 66 percent of crises are captured. A signal equal to 1 is issued when the credit-to-GDP gap exceeds the threshold. Otherwise, the signal is equal to 0. A signal of 1 at time t is judged to be correct (false) if a crisis (no crisis) occurs over the following signaling horizons: (i) "year 1 to 3": quarters $t+1$ to $t+12$, (ii) "year 1": quarters $t+1$ to $t+4$, (iii) "year 2": quarters $t+5$ to $t+8$, (iv) "year 3": quarters $t+9$ to $t+12$, (v) "year 1 and 2": quarters $t+1$ to $t+8$, and (vi) "year 2 and 3": quarters $t+5$ to $t+12$. The numbers in parentheses next to "All Data" and "Homogenous Data" correspond to the number of crises that are included in the analysis using the respective sample.

References

Aikman, D., A. G. Haldane, and B. D. Nelson. 2010. "Curbing the Credit Cycle." Speech prepared for the Columbia University Center on Capitalism and Society Annual Conference, New York, November.

Alessi, L., and C. Detken. 2011. "Quasi Real Time Early Warning Indicators for Costly Asset Price Boom/Bust Cycles: A Role for Global Liquidity." *European Journal of Political Economy* 27 (3): 520–33.

Alfaro, R., and M. Drehmann. 2009. "Macro Stress Tests and Crises: What Can We Learn?" *BIS Quarterly Review* (December): 28–41.

Angelini, A., A. Kashyap, and B. Mojon, eds. 2003. *Monetary Policy Transmission in the Euro Area.* Cambridge: Cambridge University Press.

Bank for International Settlements. 2009. "Addressing Financial System Procyclicality: A Possible Framework." BIS Note for the FSF Working Group on Market and Institutional Resilience (April). Available at http://www.financialstabilityboard.org.

Basel Committee on Banking Supervision. 2010. *Guidance for National Authorities Operating the Countercyclical Capital Buffer.* Basel, Switzerland: Bank for International Settlements. Available at http://www.bis.org.

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

Bernanke, B., M. Gertler, and S. Gilchrist. 1996. "The Financial Accelerator and the Flight to Quality." *Review of Economics and Statistics* 78 (1): 1–15.

Borio, C. 2003. "Towards a Macroprudential Framework for Financial Supervision and Regulation?" *CESifo Economic Studies* 49 (2): 181–216. Also available as BIS Working Paper No. 128 (February).

Borio, C., and M. Drehmann. 2009a. "Assessing the Risk of Banking Crises — Revisited." *BIS Quarterly Review* (March): 29–46.

———. 2009b. "Towards an Operational Framework for Financial Stability: 'Fuzzy' Measurement and Its Consequences." BIS Working Paper No. 284 (June).

Borio, C., C. Furfine, and P. Lowe. 2001. "Procyclicality of the Financial System and Financial Stability: Issues and Policy Options." In "Marrying the Macro- and Micro-Prudential Dimensions of Financial Stability." *BIS Papers* 1 (March): 1–57.

Borio, C., and P. Lowe. 2002. "Assessing the Risk of Banking Crises." *BIS Quarterly Review* (December): 43–54.

———. 2004. "Securing Sustainable Price Stability: Should Credit Come Back from the Wilderness?" BIS Working Paper No 157 (July).

Boyd, J., G. de Nicoló, and E. Loukoianova. 2009. "Banking Crises and Crisis Dating: Theory and Evidence." IMF Working Paper No. 09/141.

Brunnermeier, M., A. Crockett, C. Goodhart, A. Persaud, and H. Shin. 2009. *The Fundamental Principles of Financial Regulation.* London: Centre for Economic Policy Research. Available at http://www.cepr.org/pubs/books/P197.asp.

Bussière, M., and M. Fratzscher. 2008. "Low Probability, High Impact: Policy Making and Extreme Events." *Journal of Policy Modeling* 30 (1): 111–21.

Catarineu-Rabell, E., P. Jackson, and D. Tsomocos. 2005. "Procyclicality and the New Basel Accord — Banks' Choice of Loan Rating System." *Economic Theory* 26 (3): 537–57.

Cecchetti, S. G., M. Kohler, and C. Upper. 2009. "Financial Crises and Economic Activity." In *Financial Stability and Macroeconomic Policy.* Proceedings of the Economic Policy Symposium organized by the Federal Reserve Bank of Kansas City in Jackson Hole, Wyoming, August 20–22.

Claessens, S., M. A. Kose, and M. E. Terrones. 2011. "How Do Business and Financial Cycles Interact?" IMF Working Paper No. 11/88.

Demirgüç-Kunt, A., and E. Detragiache. 1999. "Monitoring Banking Sector Fragility: A Multivariate Logit Approach with an Application to the 1996–97 Banking Crisis." World Bank Policy Research Working Paper No. 2085.

Drehmann, M., C. Borio, L. Gambacorta, G. Jiménez, and C. Trucharte. 2010. "Countercyclical Capital Buffers: Exploring Options." BIS Working Paper No. 317.

Drehmann, M., C. Borio, and K. Tsatsaronis. 2011. "Anchoring Countercyclical Capital Buffers: The Role of Credit Aggregates." BIS Working Paper No. 355.

Drehmann, M., and L. Gambacorta. 2011. "The Effects of Countercyclical Capital Buffers on Bank Lending." Forthcoming in *Applied Economic Letters.*

Edge, R., and R. Meisenzahl. 2011. "The Unreliability of Credit-to-GDP Ratio Gaps in Real-Time and the Implications for Countercyclical Capital Buffers." Mimeo.

Gambacorta, L., and D. Marques-Ibanez. 2011. "The Bank Lending Channel: Lessons from the Crisis." *Economic Policy* 26 (66): 135–82.

Gerali, A., S. Neri, L. Sessa, and F. Signoretti. 2010. "Credit and Banking in a DSGE Model of the Euro Area." *Journal of Money, Credit, and Banking* 42 (S1): 107–41.

Gordy, M. 2009. "First, Do No Harm — A Hippocratic Approach to Procyclicality in Basel II." Paper presented at the Procyclicality in the Financial System conference, jointly organized by the Netherlands Bank and the Bretton Woods Committee, Amsterdam, February 9–10.

Gordy, M., and B. Howells. 2006. "Procyclicality in Basel II: Can We Treat the Disease without Killing the Patient?" *Journal of Financial Intermediation* 15 (3): 395–417.

Gyntelberg, J., and P. Wooldridge. 2008. "Interbank Rate Fixings during the Recent Turmoil." *BIS Quarterly Review* (March): 59–72.

Hodrick, R., and E. Prescott. 1981. "Postwar U.S. Business Cycles: An Empirical Investigation." Reprinted in *Journal of Money, Credit, and Banking* 29 (1): 1–16.

Ivashina, V., and D. Scharfstein. 2010. "Bank Lending during the Financial Crisis of 2008." *Journal of Financial Economics* 97 (3): 319–38.

Jiménez, G., and J. Saurina. 2006. "Credit Cycles, Credit Risk, and Prudential Regulation." *International Journal of Central Banking* 2 (2): 65–98.

Jordà, Ò., and A. M. Taylor. 2011. "Performance Evaluation of Zero Net-Investment Strategies." NBER Working Paper No. 17150.

Kaminsky, G., and C. Reinhart. 1999. "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems." *American Economic Review* 89 (3): 473–500.

Kashyap, A., and J. Stein. 2004. "Cyclical Implications of the Basel II Capital Standards." *Economic Perspectives* (Federal Reserve Bank of Chicago) (Q1): 18–31.

Kindleberger, C. 2000. *Manias, Panics and Crashes.* 4th ed. Cambridge: Cambridge University Press.

Kiyotaki, N., and J. Moore. 1997. "Credit Cycles." *Journal of Political Economy* 105 (2): 211–48.

Koopman, S. J., and A. Lucas. 2005. "Business and Default Cycles for Credit Risk." *Journal of Applied Econometrics* 20 (2): 311–23.

Korineck, A. 2011. "Systemic Risk-Taking: Amplification Effects, Externalities, and Regulatory Responses." Mimeo. Available at http://www.korinek.com.

Laeven, L., and F. Valencia. 2008. "Systemic Banking Crises: A New Database." IMF Working Paper No. 08/224.

———. 2010. "Resolution of Banking Crises: The Good, the Bad, and the Ugly." IMF Working Paper No. 10/146.

Lorenzoni, G. 2008. "Inefficient Credit Booms." *Review of Economic Studies* 75 (3): 809–33.

Lown, C., and D. Morgan. 2006. "The Credit Cycle and the Business Cycle: New Findings Using the Loan Officer Opinion Survey." *Journal of Money, Credit, and Banking* 38 (6): 1575–97.

Lown, C., D. Morgan, and S. Rohatgi. 2000. "Listening to Loan Officers: The Impact of Commercial Credit Standards on Lending and Output." *Economic Policy Review* (Federal Reserve Bank of New York) 6 (2): 1–16.

McGuire, P., and P. Wooldridge. 2005. "The BIS Consolidated Banking Statistics: Structure, Uses and Recent Enhancements." *BIS Quarterly Review* (September): 73–86.

Meh, C., and K. Moran. 2010. "The Role of Bank Capital in the Propagation of Shocks." *Journal of Economic Dynamics and Control* 34 (3): 555–76.

Mendoza, E. G., and M. E. Terrones. 2008. "An Anatomy of Credit Booms: Evidence from Macro Aggregates and Micro Data." NBER Working Paper No. 14049.

Minsky, H. P. 1982. "Can 'It' Happen Again?" In *Essays on Instability and Finance.* Armonk, NY: M. E. Sharpe Inc.

Orphanides, A., and S. van Norden. 2002. "The Unreliability of Output-Gap Estimates in Real Time." *Review of Economics and Statistics* 84 (4): 569–83.

Rajan, R. 2005. "Has Financial Development Made the World Riskier?" In *The Greenspan Era: Lessons for the Future.* Proceedings of the Economic Policy Symposium organized by the

Federal Reserve Bank of Kansas City in Jackson Hole, Wyoming, August 25–27.

Ravn, M., and H. Uhlig. 2002. "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations." *Review of Economics and Statistics* 84 (2): 371–76.

Reinhart, C., and K. Rogoff. 2009. *This Time Is Different: Eight Centuries of Financial Folly.* Princeton, NJ: Princeton University Press.

Repullo, R., and J. Saurina. 2011. "The Countercyclical Capital Buffer of Basel III: A Critical Assessment." Mimeo. Available at http://www.cemfi.es/∼repullo/papers.htm.

Repullo, R., and J. Suarez. 2009. "The Procyclical Effects of Bank Capital Regulation." Mimeo. Available at http://www.cemfi.es/∼repullo/papers.htm.

Schularick, M., and A. M. Taylor. Forthcoming. "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles and Financial Crises, 1870–2008." Forthcoming in *American Economic Review.*

Van den Heuvel, S. 2008. "The Welfare Cost of Bank Capital Requirements." *Journal of Monetary Economics* 55 (2): 298–320.

VanHoose, D. 2007. "Theories of Bank Behavior under Capital Regulation." *Journal of Banking and Finance* 31 (12): 3680–97.

Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton, NJ: Princeton University Press.

Zhu, H. 2008. "Capital Regulation and Banks' Financial Decisions." *International Journal of Central Banking* 4 (1): 165–211.

Discussion of "Anchoring Countercyclical Capital Buffers: The Role of Credit Aggregates"*

Òscar Jordà
University of California, Davis

1. Introduction

A tree's risk of catching fire is usually small, except when the forest is ablaze. It is hardly surprising that the international scope of the recent financial crisis has renewed efforts at designing bank capital regulation to account for both idiosyncratic and systemic risk. I say renewed because crisis prevention policies have been at the forefront of these discussions for quite some time. Almost ten years ago to the date, the Bank of England hosted a conference on "Banks and Systemic Risk" in which Howard Davies, then Chairman of the recently created Financial Services Authority in the United Kingdom, gave an overview on the use of variable capital requirements as a tool against systemic risk (see Davies 2001), the focus of the Drehmann, Borio, and Tsatsaronis paper in this issue (in the context of Basel III), which I will now discuss.

The enduring desire to endow regulation of a macroprudential orientation serves to highlight the difficulty in drawing up rules to satisfy a set of statutory objectives. The turning points and intensity of the financial cycle are not directly observable, a considerable complication for designing variable countercyclical capital buffers, whose justification needs to be specially transparent and unambiguous.

Moreover, supervisory standards should be shared globally to avoid "regulatory arbitrage." Thus, as the winds of the recent

*Author e-mail: ojorda@ucdavis.edu.

financial storm died down, the gulf between American and British views on minimum capital standards have increasingly widened. Whereas in the United Kingdom minimum requirements as high as 15–20 percent form part of the conversation, "here [meaning in the United States], that thought is restricted to cranks and university professors. . . . The banks decry any contemplation of an increase in capital standards as a death knell for economic growth" (Eisenger 2011).

Against this backdrop, it is worth remembering that the last thirty years represent an unprecedented period in the financial history of the twentieth century. A study of fourteen advanced countries over the last 140 years by Schularick and Taylor (forthcoming) demarcates three phases of financial development. From the 1870s until the 1940s, the ratio of bank loans to GDP remained very stable at around 50 percent. That stability contrasts with the numerous financial events experienced, culminating in the Great Depression. After World War II and up to the fall of Bretton Woods and the first oil crisis, these advanced economies experienced a period of unprecedented calm—no crisis events are recorded in their sample—with a moderate trend toward increased but strongly regulated financial intermediation coupled with economic growth.

What about the last thirty years? In the United States, the sum of all financial sector claims on the economy relative to annual output stood at about 150 percent in 1975. By 2008 that number had ballooned to about 350 percent (*Economic Report of the President* 2010)—far above pre–Great Depression highs. In the United Kingdom, the total balance sheet of the financial sector amounted to 34 percent of GDP in 1964, yet by 2007 it had reached 500 percent (Turner 2010). By 2008, the ratio of bank loans to GDP in the Schularick and Taylor (forthcoming) data had swelled to 200 percent from 50 percent earlier on. These numbers are coupled with the resurgence of financial crises after the post–World War II oasis of calm that culminated with the well-known events of 2007–08. Yet average growth of output from 1945 to 1975 across all fourteen countries in the Schularick and Taylor (forthcoming) paper is double the growth rate in the thirty years after 1975. In the words of Lord Turner (2010), "There is no clear evidence that the growth in the scale and complexity of the financial system in the rich developed world over the last 20 to 30 years has driven increased growth stability." And Hume and Sentance (2009) have pointed out the stagnant or falling rates of aggregate investment despite the sharp rise in outstanding credit volumes. It would be premature to conclude that the bankers' views on bank regulation are nothing short of farcical and self-serving, but prima facie evidence suggests capital requirements are more "tonic" than "death knell" for economic growth.

Schularick and Taylor (forthcoming)—and, in an extension that includes external imbalance, Jordà, Schularick, and Taylor (2011)—find that credit growth relative to GDP is the best predictor of a financial crisis, with external imbalances a natural concomitant factor in the financial events prior to World War II, but not since. And this finding meshes well with the findings of Repullo and Saurina (2011), who argue that the cyclical properties of credit growth (rather than the credit-to-GDP gap) are better suited to design a mechanical rule for countercyclical buffers: credit-to-GDP gap is negatively correlated with GDP growth in many countries, implying that a reduction of capital requirements is advised in good times, and an increase in bad times. Credit growth and GDP growth are positively correlated instead.

That external imbalances do not play a role in explaining financial crises does not mean that financial events do not propagate across borders, and in fact, the analysis in Jordà, Schularick, and Taylor (2011) finds evidence of precisely this type of phenomenon using straightforward tools from the analysis of networks. This observation provides a natural link between another strand of literature, that on financial networks (e.g., Gai and Kapadia 2010; Haldane 2009; Haldane and May 2011; May and Arinaminpathy 2010; May, Levin, and Sugihara 2008; and Nier et al. 2007), and Drehmann, Borio, and Tsatsaronis' analysis of idiosyncratic factors in connection to macroprudential indicators. May, Levin, and Sugihara (2008) liken financial systems to the type of network found in ecological systems, where highly connected "large" nodes (read large banks) tend to have their connections disproportionately with "small" nodes, and conversely, small nodes connect with disproportionately few large nodes. Such a network architecture is referred to as "disassortative."

An analysis of the U.S. Fedwire system of interbank flows commissioned by the Federal Reserve Bank of New York (Soramäki et al. 2006) showed that on a daily basis, 75 percent of payment flows involve fewer than 0.1 percent of the nodes and only 0.3 percent of observed linkages between nodes—an example of a highly disassortative network, which tends to be stable except when one of the large nodes fails. In a recent paper, Hale (2011) analyzes data for interbank loans involving nearly 8,000 banking institutions across 141 countries using network tools. Hale (2011) reports evidence that the banking network had become more dense, more clustered, and less symmetric, all of which is likely to have increased its fragility and potential for contagion. Arguably then, regulation aimed at improving the stability of the system against systemic risk would impose capital requirements as a function of the network connectivity features of each institution.

Drehmann, Borio, and Tsatsaronis provide a careful analysis of the policy options for dealing with financial system procyclicality through the use of variable bank capital requirements. While they discuss different schemes by which these requirements could be set, it is their empirical evaluation of possible financial stress indicators that is the most innovative aspect of their analysis and where I will focus my discussion. And in particular, I wish to borrow some standard protocols that are routine in medicine, meteorology, and other sciences as a way to formalize the evaluation of these stress indicators and how their signals should be used to determine appropriate capital requirements by balancing costs and benefits.

2. Financial Cycles

Ideally, a variable capital requirement rule should be as simple to communicate as Taylor's (1993) monetary policy rule is to determine how interest rates should be set in response to inflation and output deviations from target. Within the minimum capital requirement levels set by Basel II and some agreed maximum level, capital requirements could be set as a function of deviations of an observable indicator of leverage and macroprudential risk relative to an ideal target level. Requirements could be raised in response to excess leverage, and capital buffers could be depleted when leverage falls below target. Such a rule would presumably smooth the financial cycle and prevent financial crises altogether. In practice, this is much harder than it looks, since idiosyncratic and systemic factors need to be considered in setting the requirements of each institution. Moreover, what is a good leverage indicator is difficult to determine when the causes of financial events are not always well understood. And of course the rule could be generalized to include other indicators, smoothing, and asymmetries.

As Drehmann, Borio, and Tsatsaronis rightly point out, the onset and the intensity of a financial cycle are unobservable. But neither are business cycles: is a low value of GDP an indication of recession or just a temporary setback in the midst of an expansion? Yet I would argue that the same approach used to date and evaluate business-cycle chronologies can be used to analyze financial cycles. Dates of banking crises are available in Laeven and Valencia (2008) and Reinhart and Rogoff (2009), the sources of the crisis dates used by Drehmann, Borio, and Tsatsaronis. But rather than focusing on singular dates of unfortunate and rarely observed events (depending on the sample, these are events observed less than 5 percent of the time), it may be useful to consider a chronology of financial "recessions" as it were, and such is the approach in Claessens, Kose, and Terrones (2011). In that paper, credit—measured as aggregate claims on the private sector by deposit money banks—is sorted into periods of leverage upturns and downturns using the simple Bry and Boschan (1971) algorithm as refined in Harding and Pagan (2002).

The advantages of this approach are numerous. First, the data can be analyzed raw (in the levels) so there is no need to determine an adequate detrending method, of which several alternatives exist but no unique standard has been agreed upon. Second, the method is robust to the arrival of new data in the sense that the dating of the earlier part of the sample remains unaffected. The same cannot be said of most detrending methods since the trend itself will tend to vary as the sample expands (see Canova 1998). Third, the algorithm is transparent and easy to replicate and does not depend on ad hoc interpretation of the historical record.

In a sample of quarterly data from 1960:Q1 until 2007:Q4 for twenty-one advanced OECD countries, Claessens, Kose, and Terrones (2011) identify 114 periods of financial distress (about 25 percent of the time in their sample), not all of which corresponded to a financial crisis à la Laeven and Valencia (2008) and Reinhart and Rogoff (2009). Many correspond well with business-cycle turning points, but not always. On a per-country basis, financial cycles are observed slightly less frequently than business-cycle recessions. An alternative, but in my view less preferable, approach is provided in Lo Duca and Peltonen (2011).

2.1 Evaluating Chronologies of Leverage Cycles

How good are any of these chronologies and therefore which should be used as a benchmark to determine the best real-time indicators and predictors of future turning points? This task may appear hopeless, since there is no gold standard against which each candidate chronology can be compared. But at least the beginnings of an answer can be found in a recent paper by Berge and Jordà (2011), which examines this exact problem in the context of the National Bureau of Economic Research's (NBER's) dating of peaks and troughs of economic activity.

The intuition of how this can be done consists of thinking about the candidate chronology as an expression of the latent state of the financial cycle. Next, think of observable financial conditions indicators as being generated by a mixture distribution, each state determined by the candidate chronology. Sort the observed data depending on the state and compare the resulting empirical distributions. If the chronology is uninformative, a draw with a high value of the financial conditions indicator will be as likely to have come from one state as from the other. If it is informative, the opposite will be the case. Therefore, a measure of the distance between the two implied empirical distributions of the financial indicator can be used to gauge the sorting abilities of each chronology. The two standard measures available are the Kolmogorov-Smirnov statistic and the rank-sum Mann-Whitney statistic, both of which are described in Berge and Jordà (2011). Of course, one must first identify what indicator or indicators could be used for this purpose. Berge and Jordà (2011) use GDP growth and an economic activity index constructed with a factor model, and a similar approach could be used here on, say, credit growth or a financial conditions index constructed with a factor model that may include credit conditions indicators and asset prices, for example.

3. Evaluating the Anchor Variables: The Correct Classification Frontier

Drehmann, Borio, and Tsatsaronis evaluate candidate *anchor* variables—the nomenclature they use to refer to predictors of financial stress events—relative to their ability to sort when a financial event will occur. Therefore, a crisis signal is issued about an impending financial event if the anchor variables take values above a predetermined threshold. In this section, I would like to expand on two related issues: (i) a formal way to choose this threshold so as to balance the costs and benefits of each alternative, and (ii) a formal statistic of overall signaling ability. The principle at play in this section is the notion that the usefulness of a forecast should be judged by the rewards associated with the actions taken by the agent (in this case, the regulator) as a result of the forecast (see Granger and Machina 2006).

Let y_{t-h} be a candidate anchor variable available at time $t - h$ for $h = 0, 1, 2 \ldots H$, and let $S_t \in \{0, 1\}$ be an indicator of the unobserved financial state of the system, a 1 indicating a crisis (or a financial "recession" as explained earlier). At this point, it is important to remark that the Drehmann, Borio, and Tsatsaronis definition of what constitutes a successful prediction is proper classification of crisis/no-crisis periods within a three-year window. Moreover, they separate the onset of the event from its aftermath. This is not what is standard and not what I advocate here—expansions and recessions are not symmetric either—although the procedures that I discuss can be applied equally. Instead, a prediction on the state at time t can be formed as $\widehat{S}_t(h) = I(y_{t-h} > c_h)$, where $I(.)$ is the indicator function that takes the value of 1 if the argument is true, and 0 otherwise. c_h is the threshold associated with an h-periods-ahead forecast. In my view, it would be preferable to evaluate individually which variables are best for what horizons—for example, as when Berge and Jordà (2011) evaluate the information content of the components of the leading economic indicators index in the United States. For simplicity, however, I present the discussion in terms of $h = 0$ and simplify the notation to $\widehat{S}_t$. There are four possible outcomes associated with the {prediction, state} pair with the following probabilities: $P(\widehat{S}_t = 1|S_t = 1)$ is the *true positive rate* $TP(c)$; $P(\widehat{S}_t = 0|S_t = 1)$ is the *false negative rate* $FN(c)$ or type II error; $P(\widehat{S}_t = 1|S_t = 0)$ is the *false positive rate* $FP(c)$ or type I error; and $P(\widehat{S}_t = 0|S_t = 0)$ is the *true negative rate* $TN(c)$. Clearly, $TP(c) + FN(c) = 1$ and $TN(c) + FP(c) = 1$. These rates all depend on the predictive qualities of y_t and the threshold c chosen.

The decision of which threshold to choose can be analyzed using Peirce's (1884) "utility of the method," which in modern parlance characterizes the expected utility given the costs and benefits of each type of error, ability of the predictor to properly sort the data, and unconditional incidence of the phenomena under study. Specifically,

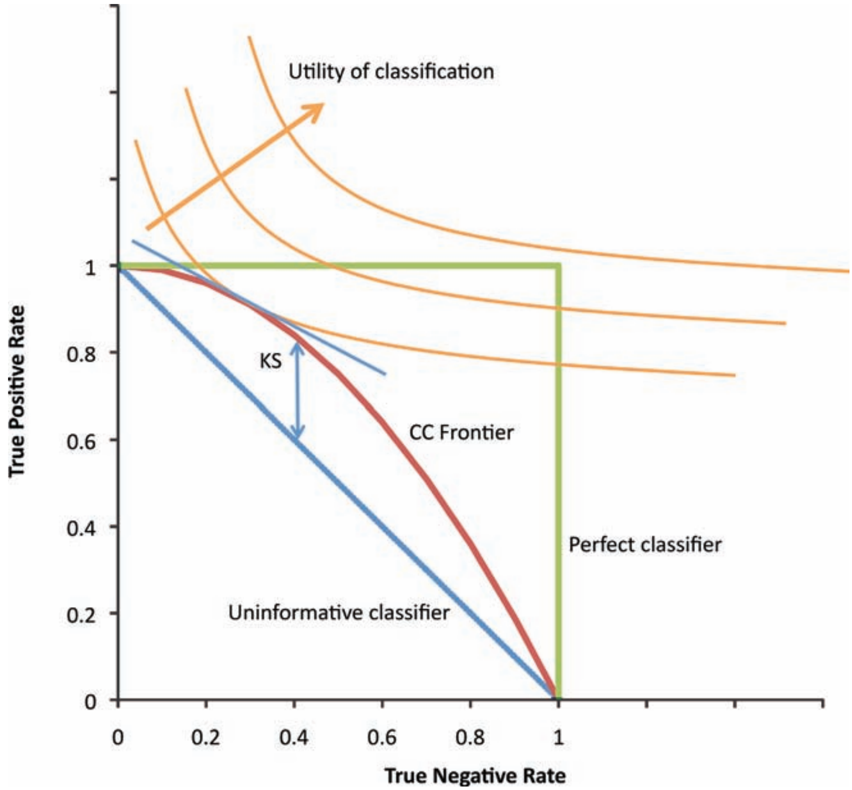
$$U(c) = U_{pP}TP(c)\pi + U_{nP}(1 - TP(c))\pi + U_{pN}(1 - TP(c))(1 - \pi) + U_{nN}TN(c)(1 - \pi), \quad (1)$$

where $\pi = P(S_t = 1)$, the unconditional probability of a financial event, and U_{aA} for $a \in \{n, p\}$ and $A \in \{N, P\}$ is the utility associated with each of the possible four outcomes defined by the {prediction, state} pair. Notice that since $TP(c) + FN(c) = 1$ and $TN(c)+FP(c) = 1$, everything can be expressed in terms of the true classification rates. Moreover, choosing $c \rightarrow \infty$ will drive $TN \rightarrow 1$ but $TP \rightarrow 0$ and vice versa. Therefore, we can plot the trade-offs in expression (1) by thinking of the combinations $\{TP(c), TN(c)\}$ for $c \in (-\infty, \infty)$ as a sort of production possibilities frontier of correct classification. This curve contains the same intuition as the production possibilities frontier for two goods in standard microeconomics, here the two "goods" being $TP(c)$ and $TN(c)$. Jordà and Taylor (2011) denominate this curve the *correct classification frontier* (CCF), and it is displayed in figure 1.

The diagonal line running from $(0, 1)$ to $(1, 0)$ is the CCF for an uninformative anchor variable since $TP(c) = 1 - TN(c)$ $\forall c$. Conversely, an anchor variable with perfect sorting abilities has a CCF that hugs the northeast corner of the $[0, 1]\times[0, 1]$ square. In that case, the relative utility weights become irrelevant because one obtains a corner solution. But in real situations, the CCF will be between these two extremes. As an example, I calculated the CCF for the real credit growth indicator in Drehmann, Borio, and Tsatsaronis, which is similar to the credit growth indicator used in Jordà, Schularick, and Taylor (2011) and Schularick and Taylor (forthcoming) with good success. This CCF is displayed in figure 2.

Further interpretability can be gained by assuming that $U_{pP} = -U_{nP}$ and $U_{nN} = -U_{pN}$; that is, think of U_{pP} as the loss of output (relative to some norm) due to a crisis, and U_{nN} as the loss of output avoided due to unnecessary capital requirements in financially calm times. In that case, the utility function can be plotted against the CCF as is done in figure 1. The optimal choice of threshold c then becomes the tangent between the CCF and the utility function,

Figure 1. The Correct Classification Frontier

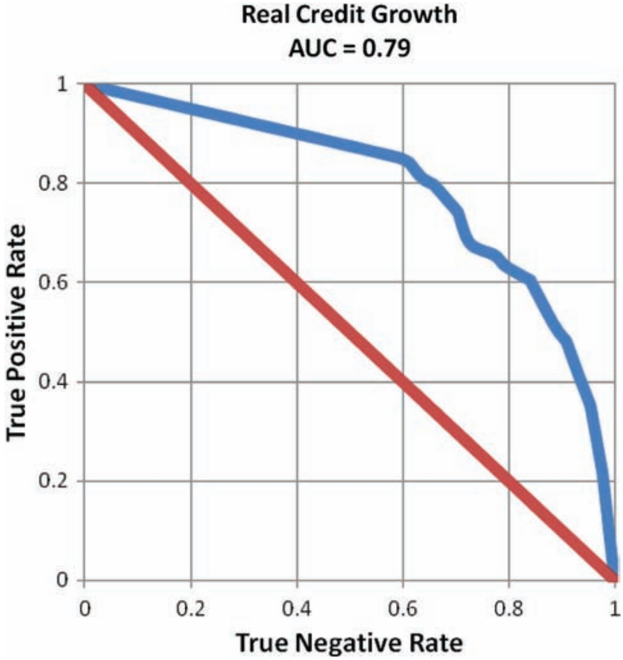


Note: KS refers to the Kolmogorov-Smirnov statistic and is the maximum distance between the CCF and the diagonal line.

which occurs when the slope of the CCF (which is the likelihood ratio of the distributions implied by the state S_t) is equivalent to the expected marginal rate of substitution between output losses due to crises and output losses due to superfluous regulation.

More texture to this analysis can be gleaned from the literature. For example, Jordà, Schularick, and Taylor (2011) calculate that, relative to non-financial crisis recessions, financial crises depress output by an additional 1 percent over a four-year window. On the other hand, Angelini et al. (2011) calculate the impact of the proposed Basel III framework to be in the order of 0.08 to 0.09 percent of

Figure 2. The Correct Classification Frontier for Real Credit Growth



Notes: AUC refers to the area under the CCF and is equal to 0.79. For an uninformative classifier, the AUC is 0.5 and for a perfect classifier, the AUC is 1. Calculated from the data in table 1 in Drehmann, Borio, and Tsatsaronis (this issue).

steady-state output per each percent increase in the requirement. If one assumes it takes nine years to converge to the new steady state, the cost of lower annual GDP growth is 0.01 percent. Clearly the first measure needs to be appropriately renormalized to be comparable to the second one, but at least these numbers can form the beginning of a conversation. And these are not the only sources. For example, the Macroeconomic Assessment Group (2010) at the Bank for International Settlements find that a 1 percent increase in capital requirements implemented over eight years causes a 0.03 percent lower annual GDP growth. The Basel Committee on Banking Supervision (2010) provides yet another set of costs and benefits to draw numbers from.

Yet, even if an agreed-upon characterization of policymaker preferences is hard to come by, the CCF can be used to compute a simple, non-parametric statistic that describes the classification ability of each anchor variable depending on the forecast horizon. This statistic is the *area under the* (CCF) curve, or AUC, and has a long tradition in biostatistics and other sciences (see Jordà and Taylor 2011 for an extensive review of the literature). This statistic can be easily calculated as

$$AUC = \frac{1}{T_N T_P} \sum_{j=1}^{T_N} \sum_{i=1}^{T_P} I\left(y_j^n < y_i^p\right),$$

where T_N and T_P refer to the number of observations associated with the two states 0 and 1, respectively (or "negatives" and "positives"), and $I(y_j^n < y_i^p)$ is just another way of counting the frequency with which values of the anchor variable in the 0 state, y_j^n, attained values that are lower than those attained when the anchor variable is in state 1, y_i^p. The AUC is a Mann-Whitney rank statistic that takes the value of 0.5 for a completely uninformative classifier (the diagonal CCF in figure 1) and the value of 1 for a perfect classifier (with CCF given by the northeast edges of the $[0,1] \times [0,1]$ square). Moreover, under standard regularity conditions (see, e.g., Hsieh and Turnbull 1996),

$$\sqrt{T}(\widehat{AUC} - 0.5) \xrightarrow{d} N(0, \sigma^2),$$

which provides the foundation to construct simple inferential procedures based on the standard Wald principle (see Jordà and Taylor 2011 for an extensive review of testing procedures and ways to estimate σ^2). In fact, most of these methods are available in commonly used econometric packages such as STATA. As an example, I summarize the output in tables 1, 2, and 3 in the Drehmann, Borio, and Tsatsaronis paper using back-of-the-envelope calculations to summarize their findings and with the caveat that they are using a three-year window to claim a successful classification. Thus, this will generate overly optimistic numbers that are reported in table 1.

Table 1 makes it easier to see that the best predictors of financial events are credit-to-GDP gap, property gap, and LIBOR OIS with values of the AUC around 0.8. This value is reasonably high: for

Table 1. The Classification Ability of Candidate Conditioning Variables Using the AUC

Conditioning Variable	AUC	Conditioning Variable	AUC
Credit-to-GDP Gap	0.821	Profits Before Tax	0.712
Real Credit Growth	0.756	Bank CDS	0.535
Real GDP Growth	0.672	Credit Spread	0.730
Equity Gap	0.705	Credit Spread Long T.	0.594
Property Gap	0.821	LIBOR OIS	0.798
NPL	0.584	TED Spread	0.664

Notes: AUC refers to the area under the correct classification frontier. An uninformative classifier has a value of 0.5, whereas a perfect classifier attains a value of 1. The numbers in the table are calculated using the data in tables 1, 2, 3, and 4, in Drehmann, Borio, and Tsatsaronis (this issue).

example, the widely used prostate-specific antigen (PSA) blood test has a similar AUC for detecting prostate cancer (see, e.g., Thompson et al. 2005), and the S&P 500 has an AUC of 0.86 for detecting in current time whether the economy is in recession or not. Similarly, non-performing loans, bank credit default swaps, and the long-term credit spread emerge as the worst indicators with AUC values that, had they not been calculated over three-year windows, would probably be no better than a coin toss.

It is well known that the best predictions are often found by combining the information of all the available predictors, and in this case, this can be easily accomplished by specifying the log-odds ratio of a financial event as a linear combination of anchor variables. Berge and Jordà (2011) use precisely this approach to determine the optimal linear combination of leading economic indicators and show that this linear combination depends heavily on the forecast horizon. This suggests that the de facto aggregation into three-year windows in the Drehmann, Borio, and Tsatsaronis paper may in fact be obfuscating a more clear determination of what anchor variables work and over what horizons.

One final note about the construction of macroeconomic gaps as deviations from a Hodrick and Prescott (HP) trend is worth making. Any monotone transformation of the anchor variable y_t delivers the same CC frontier and hence the same AUC. Unfortunately, the manner in which the trend is calculated, and hence y_t is constructed, does not fall into this category: construction of macroeconomic gaps using different trends will result in different classification ability (or different AUC values). It may well be that there is little difference between HP trends constructed with different values of λ (as the robustness analysis in the Drehmann, Borio, and Tsatsaronis paper indicates), and a value of $\lambda = 400{,}000$ certainly makes the HP trend virtually equivalent to calculating a linear trend. But there are other reasons to find the approach wanting.

To facilitate intuition, think of linear detrending, where the trend is recalculated each time a new observation arrives. Depending on the unknown data-generating process, the most likely outcome is that trend estimates will vary with the sample size. And this will cause gaps calculated earlier to vary, making difficult any retrospective evaluation. Moreover, and as I indicated in the introduction, Repullo and Saurina (2011) suggest that the credit-to-GDP gap moves countercyclically (with output) in many countries, which would seem to suggest contra naturam that capital standards ought to be raised in bad times and lowered in good times. On the other hand, as Repullo and Saurina (2011) point out, credit growth is procyclical and has the advantage of being easy to communicate and stable over time, a desirable feature in a regulatory environment where the regulator is likely to be closely scrutinized.

4. Financial Network Connectivity and Event Detection

My last methodological comment presents a simple way to include the network connectivity properties of individual banks as a way to adjust capital requirements to that bank's systemic relevance within the financial network. There is a substantial but recent literature based on network analysis that suggests that a network's stability depends, in an important and highly non-linear manner, on its topology. Financial networks evolve endogenously, usually with a few large but superconnected clusters of nodes relative to a vast cluster of small but sparsely connected nodes—in the network parlance, a "disassortative" topology. May and Arinaminpathy (2010) explain that such an architecture can endow a network of good stability properties except when one or more of the large nodes gets knocked out. Interestingly, they find the ratio of capital buffers to assets to be relatively smaller in bigger banks than in smaller banks, exactly the reverse of what would seem preferable. This point has been emphasized in a recent speech by Haldane (2009).

Here I would like to offer one approach to constructing and evaluating the virtues of a network connectivity indicator for the purpose of predicting impending financial events and possibly as a variable to be used to determine variable requirements for banks of different import. Jordà, Schularick, and Taylor (2011) use such an approach to examine the international contagion properties of financial crises and find that there is predictive value in knowing if other countries have experienced a crisis in the recent past.

I begin by proposing a way to evaluate how network connectivity of node-level financial events can be used to predict systemic events. Two straightforward network connectivity measures are the *incidence rate*, r_t, and the *wiring ratio*, w_t. In this application, the incidence rate simply measures the proportion of banks experiencing a financial event (where a financial event at the bank level needs to be more precisely defined—but here let's suppose for simplicity, a bank failure) relative to the totality of banks. Thus, let $S_{it} \in \{0, 1\}$, with 1 indicating that bank i has experienced a financial event at time t. Hence define

$$\eta_t = \sum_{i=1}^{n} S_{it}$$

so that $r_t = \eta_t / n$, n being the total number of banks. There are three unsatisfactory features of such a simple measure: (i) it does not account for the size of the institution, (ii) the marginal effect of an additional bank failure is independent of how many banks have failed, and (iii) it does not account for the nature of its pairwise connections to other institutions.

The wiring ratio, weighted by the size of the pairwise connections to other institutions relative to the system, offers a remedy to these three shortcomings. Define the ratio of pairwise connections between "failing" banks to total pairwise connections as

$$w_t = \frac{\eta_t(\eta_t - 1)}{n(n-1)} \in [0, 1]. \quad (2)$$

Before we bring in the weighting, it is clear that the marginal effect of an additional bank failure is quite low when few banks are in distress, but quite high in the alternative. In order to introduce weighting, consider a simple example and thus define g_{it} as some measure of the banks' assets, although some other measure could be used instead. Therefore, a natural weighted version of (2) is

$$w_t^* = \frac{\mathbf{S}_t' G_t \mathbf{S}_t}{(n-1)\sum_{i=1}^{n} g_{it}} \in [0,1], \tag{3}$$

where $\mathbf{S}_t$ is an $n \times 1$ vector with 1's in entries for banks experiencing a financial event, and 0 otherwise. G_t is an $n \times n$ lower triangular matrix with zeroes in the main diagonal and with typical (i,j) entry given by $g_{it} + g_{jt}$ for $i > j$. Thus, the weighted wiring ratio computes the ratio of pairwise weighted connections of banks in distress relative to the total sum of bank-size weighted pairwise connections. Of course, here I am focusing on measures of financial distress connectivity, but one could choose instead connectivity on the basis of interbank flows, for example. The ability of w_{t-h}^* to determine a financial event at time t can then be evaluated using the techniques described in the previous section.

Focusing instead on a bank's connectivity (say, by measuring the flows of interbank payments), a different measure of network connectivity could be constructed as follows. Define Γ_t as an $n \times n$ matrix with typical entry $\gamma_t(i,j)$ representing the payments of bank i to bank j and with $\gamma_t(i,i) = 0$ by construction. Let L_{it} be the i^{th} row of matrix Γ_t, which collects all the entries in which bank i makes a payment to other banks (L is for *liabilities*). Conversely, let A_{it} be the i^{th} column of the matrix Γ_t, which collects all the payments made to bank i (A is for *assets*). Then, the ratio of total payments made and received by bank i relative to the total volume of pairwise payments in the system is

$$\omega_{it}^* = \frac{L_{it}\mathbf{1}_n + \mathbf{1}_n' A_{it}}{\mathbf{1}_n' \Gamma_t \mathbf{1}_n} \in [0,1]$$

and can be seen as a bank-specific wiring ratio analogue to (3), but based on payment flows rather than on defaults.

The measures proposed here are only meant to be illustrative of how network connectivity can be characterized for the purposes of predicting financial events and for the purposes of tailoring capital requirements to institutions so as to account for systemic risk.

Clearly, more work is required in this particular area, although there are some interesting recent estimates in Schwaab, Koopman, and Lucas (2011), who analyze 12,000 firms and extract an estimate of what they label a default risk factor.

5. Summary

Ideally, one would formulate a dynamic stochastic general equilibrium model of the economy with a detailed characterization of how a heterogeneous financial sector affects real activity, how systemic risk can build up and wane, how capital regulation interacts with economic growth, and whether it can be crafted to smooth the financial cycle and prevent financial events. But we are far from having such an understanding and in the meantime the design of a supervisory framework that can meet real and pressing needs must rely primarily on a careful statistical analysis of the data. Drehmann, Borio, and Tsatsaronis have made a decisive contribution in this direction to which I only wish to add a bit of structure to think about the problem more formally.

References

Angelini, P., L. Clerc, V. Cúrdia, L. Gambacorta, A. Gerali, A. Locarno, R. Motto, W. Roeger, S. Van den Heuvel, and J. Vlček. 2011. "BASEL III: Long-Term Impact on Economic Performance and Fluctuations." Federal Reserve Bank of New York. Staff Report No. 485.

Basel Committee on Banking Supervision. 2010. "An Assessment of the Long-Term Economic Impact of Stronger Capital and Liquidity Requirements." Bank for International Settlements. Available at http://www.bis.org/publ/bcbs173.pdf.

Berge, T. J., and Ò. Jordà. 2011. "Evaluating the Classification of Economic Activity into Recessions and Expansions." *American Economic Journal: Macroeconomics* 3 (2): 246–77.

Bry, G., and C. Boschan. 1971. "Cyclical Analysis of Economic Time Series: Selected Procedures and Computer Programs." NBER Technical Paper No. 20.

Canova, F. 1998. “Detrending and Business Cycle Facts, A User’s Guide.” *Journal of Monetary Economics* 41 (3): 533–40.

Claessens, S., M. A. Kose, and M. E. Terrones. 2011. “Financial Cycles: What? How? When?” IMF Working Paper No. 11/76.

Davies, H. 2001. “Capital Requirements and Crisis Prevention Policies.” Remarks prepared for the Banks and Systemic Risk Conference, Bank of England, May 25. Available at http://www.bankofengland.co.uk/publications/events/conf0105/paper12may01.pdf.

Economic Report of the President. 2010. Washington, DC: U.S. Government Printing Office.

Eisenger, J. 2011. “In Debate Over Bank Capital Rules, a Trans-Atlantic Gulf.” *New York Times*, March 31, p. B4.

Gai, P., and S. Kapadia. 2010. “Contagion in Financial Networks.” *Proceedings of the Royal Society A* 466: 2401–23.

Granger, C. W. J., and M. J. Machina. 2006. “Forecasting and Decision Theory.” In *Handbook of Economic Forecasting*, Vol. I, ed. G. Elliott, C. W. J. Granger, and A. Timmermann, 81–98. Amsterdam, The Netherlands: Elsevier.

Haldane, A. G. 2009. “Rethinking the Financial Network.” Speech delivered at the Financial Student Association, Amsterdam. Available at http://www.bankofengland.co.uk/publications/speeches/2009/speech386.pdf.

Haldane, A. G., and R. M. May. 2011. “Systemic Risk in Banking Ecosystems.” *Nature* 469: 351–55.

Hale, G. 2011. “Bank Relationships, Business Cycles, and Financial Crises.” Mimeo, Federal Reserve Bank of San Francisco.

Harding, D., and A. Pagan. 2002. “Dissecting the Cycle: A Methodological Investigation.” *Journal of Monetary Economics* 49 (2): 365–81.

Hsieh, F., and B. W. Turnbull. 1996. “Nonparametric and Semiparametric Estimation of the Receiver Operating Characteristics Curve.” *Annals of Statistics* 24 (1): 25–40.

Hume, M., and A. Sentance. 2009. “The Global Credit Boom: Challenges for Macroeconomics and Policy.” *Journal of International Money and Finance* 28 (8): 1426–61.

Jordà, Ò., M. Schularick, and A. M. Taylor. 2011. “Financial Crises, Credit Booms, and External Imbalances: 140 Years of Lessons.” *IMF Economic Review* 59 (2): 340–78.

Jordà, Ò., and A. M. Taylor. 2011. "Performance Evaluation of Zero Net-Investment Strategies." NBER Working Paper No. 17150.

Laeven, L., and F. Valencia. 2008. "Systemic Banking Crises: A New Database." IMF Working Paper No. 08/224 (November).

Lo Duca, M., and T. A. Peltonen. 2011. "Macro-Financial Vulnerabilities and Future Financial Stress: Assessing and Predicting Systemic Events." ECB Working Paper No. 1311.

Macroeconomic Assessment Group. 2010. "Assessing the Macroeconomic Impact of the Transition to Stronger Capital and Liquidity Requirements—Final Report." Bank for International Settlements. Available at http://www.bis.org/publ/othp12.htm.

May, R. M., and N. Arinaminpathy. 2010. "Systemic Risk: The Dynamics of Model Banking Systems." *Journal of the Royal Society Interface* 7 (46): 823–38.

May, R. M., S. A. Levin, and G. Sugihara. 2008. "Complex Systems: Ecology for Bankers." *Nature* 451: 893–95.

Nier, E., J. Yang, T. Yorulmazer, and A. Alertorn. 2007. "Network Models and Financial Stability." *Journal of Economic Dynamics and Control* 31 (6): 2033–60.

Peirce, C. S. 1884. "The Numerical Measure of the Success of Predictions." *Science* 4 (93): 453–54.

Reinhart, C. M., and K. S. Rogoff. 2009. *This Time is Different: Eight Centuries of Financial Folly.* Princeton, NJ: Princeton University Press.

Repullo, R., and J. Saurina. 2011. "The Countercyclical Capital Buffer of Basel III: A Critical Assessment." Mimeo, Banco de España.

Schularick, M., and A. M. Taylor. Forthcoming. "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles and Financial Crises, 1870–2008." Forthcoming in *American Economic Review.*

Schwaab, B., S. J. Koopman, and A. Lucas. 2011. "Systemic Risk Diagnostics: Coincident and Early Warning Signals." ECB Working Paper No. 1327.

Soramäki, K., M. L. Bech, J. Arnold, R. J. Glass, and W. E. Beyeler. 2006. "The Topology of Interbank Payment Flows." Federal Reserve Bank of New York Staff Report No. 243.

Taylor, J. B. 1993. "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy* 39: 195–214.

Thompson, I. M., D. P. Ankerst, C. Chi, M. S. Lucia, P. J. Goodman, J. J. Crowley, H. L. Parnes, and C. A. Coltman. 2005. "Operating Characteristics of Prostate-Specific Antigen in Men with an Initial PSA Level of 3 ng/ml or Lower." *Journal of the American Medical Association* 294 (1): 66–70.

Turner, A. 2010. "What Do Banks Do, What Should They Do and What Public Policies Are Needed to Ensure Best Results for the Real Economy?" Speech given at Cass Business School, March 17.

The Unreliability of Credit-to-GDP Ratio Gaps in Real Time: Implications for Countercyclical Capital Buffers*

Rochelle M. Edge[a] and Ralf R. Meisenzahl[b]
[a]Office of Financial Stability Policy and Research, Federal Reserve Board
[b]Division of Research and Statistics, Federal Reserve Board

Macroeconomists have long recognized that activity-gap measures are unreliable in real time and that this can present serious difficulties for stabilization policy. This paper investigates whether the credit-to-GDP ratio gap, which has been proposed as a reference point for accumulating countercyclical capital buffers, is subject to similar problems. We find that ex post revisions to the U.S. credit-to-GDP ratio gap are sizable and as large as the gap itself, and that the main source of these revisions stems from the unreliability of end-of-sample estimates of the series' trend rather than from revised estimates of the underlying data. The paper considers the potential costs of gap mismeasurement. We find that the volume of lending that may incorrectly be curtailed is potentially large, although loan interest rates appear to increase only modestly.

JEL Codes: E61, G28.

 We thank Jeremy Rudd for useful comments on several of the filtering issues considered in this paper and Hilde Bjørnland, Iman van Lelyveld, and Simon van Norden for helpful discussant comments. We also thank conference and seminar participants at the Bank of England, Bank of Mexico, Norges Bank, Reserve Bank of Australia, and Reserve Bank of New Zealand and Michelle Welch for excellent research assistance. The views expressed here are our own and do not necessarily reflect the views of the Board of Governors or the staff of the Federal Reserve System. Author e-mails: rochelle.m.edge@frb.gov and ralf.r.meisenzahl@frb.gov. Mailing address: Federal Reserve Board, 20th and C Streets, NW, Washington, DC 20551.

1. Introduction

Macroeconomists have long recognized that real-time activity-gap measures are unreliable and that this can present serious difficulties for economic stabilization policies (see Orphanides and van Norden 2002 and Staiger, Stock, and Watson 1997). The use of the nominal credit-to-GDP ratio gap as a reference point for determining the need for accumulating countercyclical capital buffers—as described in the countercyclical capital buffer proposal of the Basel Committee's Macro Variables Task Force (MVTF)—could potentially suffer from the same problem, thus making it difficult to implement this type of macroprudential policy. This paper follows the approach employed by Orphanides and van Norden to investigate the relevance of this issue for the United States. Specifically, we calculate estimates of the credit-to-GDP ratio gap that would have been obtained in real time and that therefore would have been used in practice in setting countercyclical buffers. We then compare these real-time estimates with ex post estimates based on all information. Finally, we estimate the potential costs of gap mismeasurement in a practical policy context and find that the volume of lending that could be incorrectly curtailed is large.

As noted by Orphanides and van Norden, there are several reasons why real-time estimates of gap measures can differ from their final estimates. First, the underlying data used to calculate the gap measures can be revised. Second, as data in later periods become available, they could alter our estimate of where the *trend* credit-to-GDP ratio—and therefore the gap—was in earlier periods. Third, incoming data may cause us to revise our model of the time series of the credit-to-GDP ratio that we use to estimate the trend and gap. We consider all three sources of revisions to the U.S. credit-to-GDP ratio gap, although we only consider revisions that arise from the first source (changes to the data) for a limited period of time.[1]

The MVTF's consultative document considers only one method—the Hodrick-Prescott (HP) filter—for extracting the trend

[1] In particular, electronically stored vintages of data are only available starting in 1995:Q2. Earlier vintages of data are available in hard-copy form; however, because we found that data revisions contributed only modestly to total gap revisions, we decided against extending our real-time data set to earlier periods.

credit-to-GDP ratio from the actual ratio. While this is a standard method for trend extraction, it is by no means the only one. We therefore consider a range of detrending methods, albeit with the focus on whether they have different revision properties.[2]

We find that revisions to the U.S. credit-to-GDP ratio gap are sizable and are on the same order of magnitude as the gap itself. Moreover, the correlation between gaps estimated in real time and gaps estimated based on all available data are low. The main source of the revision is not from revised estimates of the underlying data but rather from the unreliability of end-of-sample estimates of the ratio's trend. Some of the episodes in which we find large revisions to the estimated gap correspond to periods in which the real-time estimate of the gap would have suggested a deployment of countercyclical capital buffers but the final gap would not. We focus on two periods where this was the case, 2001:Q4 and 2003:Q2, and calculate the potential cost of a policy implemented in real time. Using the capital ratios of U.S. banks in the Reports of Conditions and Income, we calculate the change in the systemwide capital shortfall or surplus implied by deploying the capital buffers and derive implications for lending and loan rates. We find that this policy would have acted as an additional drag on the U.S. economy following the 2001 recession.

Although the MVTF's consultative document does not consider the real-time reliability of gap measures, its analysis is mindful of the distinction between gaps estimated using data available through a particular date and gaps estimated using all available data. For example, the credit-to-GDP ratio gaps reported in the document all extract the *one-sided* Hodrick-Prescott trend, because this measure only uses information that is available at the time of the observation period.[3] That said, the document identifies "protecting the banking sector from periods of excess credit growth" as the objective of countercyclical capital buffers (p. 1) and argues that excessive credit growth is well captured by sizable deviations of the credit-to-GDP

[2]Clearly, different methods of trend extraction giving conflicting signals is another source of uncertainty in considering whether credit levels are excessive. This complication is a different issue from what is our primary concern in this paper and so we do not dwell on it further.

[3]Other papers on the detection of asset price booms, such as Alessi and Detken (2009) and Borgy, Clerc, and Renne (2009), also make this this distinction.

ratio above trend (p. 18). If we accept that the credit-to-GDP ratio gap provides a good gauge of excess credit growth, then obtaining an accurate measure of the gap is important. Of course, the best estimate of the ratio's trend—and therefore the gap—is obtained from using all available data. But it is exactly this "final" version of the gap that is poorly captured by real-time estimates, according to our analysis.

This paper is organized as follows. Section 2 outlines the detrending methods that we consider and section 3 describes the data that we employ and the "real-time" and "final" trend-estimate concepts that we use. (These concepts—and our taxonomy—are identical to those laid out by Orphanides and van Norden.) Section 4 reports our results, including our credit-to-GDP ratio gap estimates, the magnitudes of our gap revisions, and some of the effects these revisions can have in real time and ex post on policy decisions. Section 5 discusses these results, while section 6 gauges the potential costs of gap mismeasurement in terms of the volume of lending that could be incorrectly curtailed (or increase in interest rates that could obtain) as a result of basing policy decisions on misleading real-time estimates of the credit-to-GDP ratio gap. Finally, section 7 concludes.

2. Detrending Methods

The credit-to-GDP ratio that we consider exhibits distinct upward drift, likely reflecting financial deepening over time. Thus, policymakers would want to consider deviations from the upwardly trending path of the credit-to-GDP ratio in deciding whether to require banks to accumulate countercyclical capital buffers.

We use a number of different detrending methods to estimate the credit-to-GDP ratio gap. All detrending methods separate a series c_t into a trend component μ_t and a cyclical component z_t; that is, $c_t = \mu_t + z_t$. Some methods assume that the trend is a deterministic function of time; examples include linear trends, quadratic trends, cubic trends, and cubic splines. Another method of detrending a series involves applying the HP filter to the series; this is the trend-extraction method used in the MVTF's consultative document. The final approach that we consider is that of frequency detrending methods. These methods view economic time series as being the weighted sum of periodic functions (sines and cosines) and consider a series' trend to be that portion of the series that is accounted for by functions that fall within a specified frequency range. We implement this method using an approximate band-pass filter. Our discussion of these various detrending methods is brief since these methods are well described elsewhere in other sources—see, for example, Canova (1998) and Orphanides and van Norden (2002).[4]

2.1 Deterministic Detrending Methods

Deterministic detrending methods assume that the trend credit-to-GDP ratio can be well approximated by a polynomial function of time, such as a linear, quadratic, or cubic trend. The trend is the predicted value from a least-squares regression of the credit-to-GDP ratio on a constant and a scaled polynomial function of time; the gap term is the residual from the estimated equation.

If structural change is present, a single deterministic process may not be appropriate for modeling the trend of a series over the entire sample period. We therefore also use a cubic spline detrending procedure in which we divide our sample period into three equal subperiods and fit a separate cubic polynomial over each. At the boundary between two subperiods (called a "knot point"), we restrict the two spline segments to have equal values and equal first and second derivatives (this is essentially a smoothing restriction). Note, our knot points will move when we undertake our real-time analysis.

2.2 The Hodrick-Prescott Filter

The HP filter (Hodrick and Prescott 1980) optimally extracts a smooth stochastic trend that is uncorrelated with the residual

[4]In addition to the methods discussed in the paper, we also considered several unobserved-components (UC) model detrending approaches and the Beveridge-Nelson procedure. We ultimately decided not to use the UC procedures because they were very sensitive to starting values, thus making their use in a real-time exercise with different vintages or sample periods quite impractical. Moreover, the trends implied by these approaches appeared reasonably close to what we obtained with a simple linear trend. In addition, we decided not to use the Beveridge-Nelson procedure since it implied gap estimates that seemed obviously implausible.

cyclical component. Specifically, the trend credit-to-GDP ratio is derived from the HP-filter optimization problem:

$$\min_{\{\mu_t\}_{t=0}^{T}} \sum_{t=0}^{T}(c_t - \mu_t)^2 + \lambda \sum_{t=1}^{T-1}((\mu_{t+1} - \mu_t) - (\mu_t - \mu_{t-1}))^2,$$

where the value of the parameter λ governs the smoothness of the trend. When output is the variable being filtered, λ is typically set at 1,600 since this implies a business-cycle frequency of around $7\frac{1}{2}$ years. For the credit-to-GDP ratio, the MVTF's consultative document considered a range of values for the smoothing parameter; namely, $\lambda = 1{,}600 = 1^4 \cdot 1{,}600$; $\lambda = 25{,}000 \approx 2^4 \cdot 1{,}600$; $\lambda = 125{,}000 \approx 3^4 \cdot 1{,}600$; and $\lambda = 400{,}000 \approx 4^4 \cdot 1{,}600$, which are equivalent to the credit cycle's being the same, double, triple, and quadruple the length of the business cycle. Note that as $\lambda \to \infty$, the process for $\{\mu_t\}_{t=0}^{T}$ approaches a linear trend.

2.3 Frequency Detrending Methods

Frequency detrending methods model a time series as a weighted sum of periodic functions (cosines and sines) whose frequencies range from 0 to π. Cycles are fluctuations within a specified range of periodicities or frequencies (where frequency = $2 \cdot \pi$/periodicity).

As documented by King and Rebelo (1993), the HP filter is an approximation to a frequency-based "high-pass" filter that passes though the higher-frequency fluctuations in a series to the cycle while removing the low-frequency (that is, trend) fluctuations. The HP filter with $\lambda = 1{,}600$ approximates a high-pass filter that associates the cyclical component of a series with periodicities that range up to thirty-two quarters in length (and frequencies that exceed $2 \cdot \pi/32$). Another class of frequency-based filters are "band-pass" filters, which pass through higher-frequency fluctuations, only up to a specified point. That is, certain high-frequency fluctuations (such as fluctuations that might reflect residual seasonality in the data) are also excluded from the cyclical component along with the low-frequency (trend) components. When we consider band-pass filters, we follow the MVTF's approach of using a range of periodicities to represent the credit cycle. We first examined periodicities in the band of six to thirty quarters ($7\frac{1}{2}$ years), which imply frequencies ranging from $\frac{\pi}{15}$ to $\frac{\pi}{3}$ and which result in a credit cycle of about the same length as the business cycle. We also considered periodicities in the band of six to sixty quarters (15 years), which imply frequencies in the band of $\frac{\pi}{30}$ to $\frac{\pi}{3}$, and six to ninety quarters ($22\frac{1}{2}$ years), which imply frequencies in the band of $\frac{\pi}{45}$ to $\frac{\pi}{3}$. These yield credit cycles that are two and three times the length of the business cycle, respectively.

An exact band-pass filter cannot be implemented in practice since it requires a two-sided infinite-order moving average. We use Baxter and King's (1999) finite moving-average approximation, where, following Staiger, Stock, and Watson (2001), we set the width of the two-sided moving average to 160 quarters to reflect the high degree of persistence in the series we are detrending.

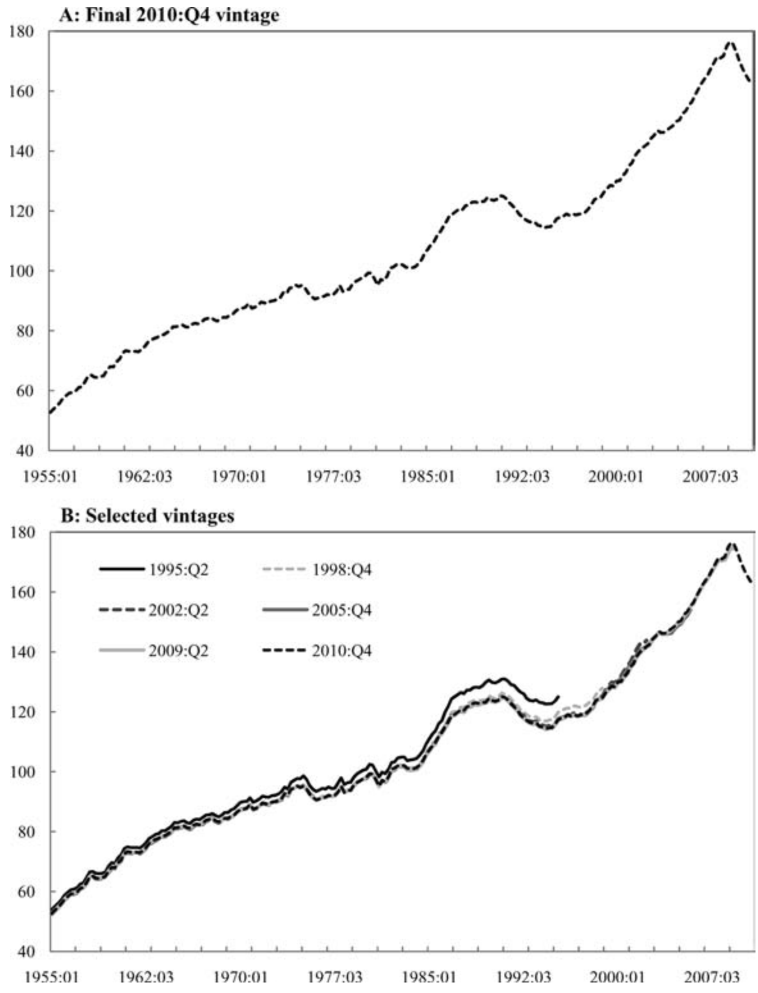
3. Data and Real-Time/Final Gap Concepts

Before reviewing the detrending methods that we use in our study, it is useful to take a look at the series that we are trying to separate into its trend and cycle (gap). The black line in panel A of figure 1 reports the time series of the nominal credit-to-GDP ratio using the definitions of variables given in the MVTF's consultative document and based on 2010:Q4 vintage data. For nominal credit (in the numerator), this is the volume of credit market debt outstanding of the non-financial corporate business sector and household and non-profit organization sector as reported by the Federal Reserve Board (FRB) in the Flow of Funds Accounts (FOFAs). For nominal GDP (in the denominator), this is the measure reported by the Bureau of Economic Analysis (BEA) in the National Income and Product Accounts (NIPAs).

3.1 Real-Time Data

Statistical agencies revise data as a result of new source data becoming available over time, reestimation of seasonal factors, and—most comprehensively—changes in series definitions. These revisions can affect the most recent quarter of the data, the last few years, or the entire history of the series. Revisions to the underlying data used to calculate credit-to-GDP ratios are one reason why gap estimates can change. Studying this source of revision requires that

Figure 1. Credit-to-GDP Ratios



we obtain real-time *time series* for our nominal credit and nominal GDP series. We obtained our real-time time series for nominal GDP from "ALFRED," which is a data archive maintained by the Federal Reserve Bank of St. Louis. Only a few real-time vintages for nominal credit are available in the ALFRED archive, however. We obtained vintages for these time series from Federal Reserve sources, although past vintages of these time series are only available in electronic form from 1995:Q2.

Although the credit and GDP series are both quarterly, the timing of their releases relative to the data's reference quarter are different. For GDP (and the NIPAs more generally) there are three estimates that are released for a given quarter: the first release, released one month after the reference period; the second release, released two months after; and the third release, released three months after. After the third release, the data are typically revised only in annual or comprehensive revisions. With credit (and the FOFAs) there is only one release of the data that usually occurs about two-and-a-half months after the reference period. Like the NIPAs, there are annual revisions to the FOFAs as well as revisions that occur following NIPAs revisions. In calculating the nominal credit-to-GDP ratio in real time, we use for nominal GDP only the time-series vintages corresponding to the third release; given the different timings of the GDP and credit data releases, this seems to be the most likely value that policymakers would use.

A vintage of data corresponds to the entire time series of the data at a particular point in time. Panel B of figure 1 plots five vintages of the nominal credit-to-GDP ratio. The name assigned to each vintage corresponds to its last observation. As can seen from panel B, each vintage plotted is different, though the differences are not large. Thus, based on observing the various vintages of the credit-to-GDP ratio, one might be inclined to conclude that real-time measurement issues of the credit-to-GDP ratio gap are unlikely to be significant. Our results demonstrate that this is not the case: Real-time measures of the credit-to-GDP ratio gap revise substantially, although—consistent with what is evident from figure 1—the major source of revision is not revisions to the data.

3.2 Estimates of the Nominal Credit-to-GDP Ratio Gap

3.2.1 True Real-Time Estimates

We have sixty-three different vintages of the nominal credit-to-GDP ratio. To obtain the true real-time estimates of the nominal credit-to-GDP ratio gap, we first apply our filtering methods (described in section 1) to each of the sixty-three credit-to-GDP ratio vintages, thereby calculating a time series of gaps for each vintage. We then take the last observation of each time series and combine it into a single series.

3.2.2 True Final Estimates

The true final estimate of the gap uses the full sample of the most recent available vintage of data. For the deterministic trends, the polynomial functions that define the trend are estimated using the most recent vintage of credit-to-GDP ratio data. There are no parameters to be estimated for the HP filter; in this case, the true final estimate of the gap is obtained from the two-sided estimate of the trend. There are also no parameters to be estimated when frequency detrending methods are employed; the true final estimate of the gap here uses—where available—subsequent periods' observations in the moving-average calculations that yield the cyclical component of the credit-to-GDP ratio in a given period.

3.2.3 Quasi Real-Time Estimates

As noted in the introduction, there are several reasons why the true real-time estimates of the gap should differ from the final estimates; only one of these involves revisions to the data in real time. In order to gauge the role of data revisions alone, we calculate quasi real-time estimates of the gap; these are calculated in a similar way to the true real-time gaps, but instead of applying our filtering methods to each of our sixty-three time-series vintages, we instead apply our filtering methods to the data from our final data vintage, with a rolling endpoint that is set equal to the period for which the gap is being calculated. We then take the last observation of each time series (as we did for the true real-time gap) and combine it into a single series. Thus, the estimate of the gap in any period only uses data up to that point in time, although we use the most recent vintage of data (not the vintage actually available at that time).[5]

[5]Orphanides and van Norden consider a fourth gap-estimate concept, called the quasi final estimate. This type of gap estimate is only relevant for gaps from unobserved-components models, which we do not consider in this paper.

4. Results

Panel A of figure 2 plots the true real-time estimates of our credit-to-GDP ratio gap over the period for which we have real-time data, with the gaps measured using all of the detrending methods described in section 1. Panel B plots the true final estimates, and panel C plots the quasi real-time estimates; for these cases we report results over the 1980:Q1 to 2010:Q4 sample period. With the exception of the cubic spline, the different methods for estimating the trend of the credit-to-GDP ratio yield gaps that display similar contours. The magnitudes of the gaps are quite different, reflecting the fact that for some methods the filter's parameters were set to pass through a greater or lesser share of the credit-to-GDP ratio's fluctuations to the cyclical component.

We begin this section by considering the sources of revision to the real-time credit-to-GDP ratio gap and then examining the magnitudes of the revisions. We then investigate the extent to which revisions could result in different policy actions being taken. After this we examine how quickly revisions tend to be realized, which is a question that could be a concern given that banks have a year in which to build their capital buffers once policymakers call for their deployment. Finally, we discuss how real-time revisions differ across filtering methods.

4.1 Revision Sources

The six panels in figure 3 plot revisions to the credit-to-GDP ratio gap. The three upper panels plot the total revision—that is, the difference between the real-time estimate and the final estimate—while the three lower panels plot the difference between the quasi real-time estimate and the final estimate. Each column in the figure corresponds to a different filtering method.

Comparing the real-time to final and quasi real-time to final revisions reveals that the revisions are fairly similar in terms of both size and contour. Recall that the reason why we constructed the quasi real-time estimate was that revisions to the underlying data represent only one possible source of revisions to the estimated gap. Even without revisions to past data, the availability of data for later periods can alter our estimate of where the trend credit-to-GDP ratio

Figure 2. Credit-to-GDP Ratio Gaps

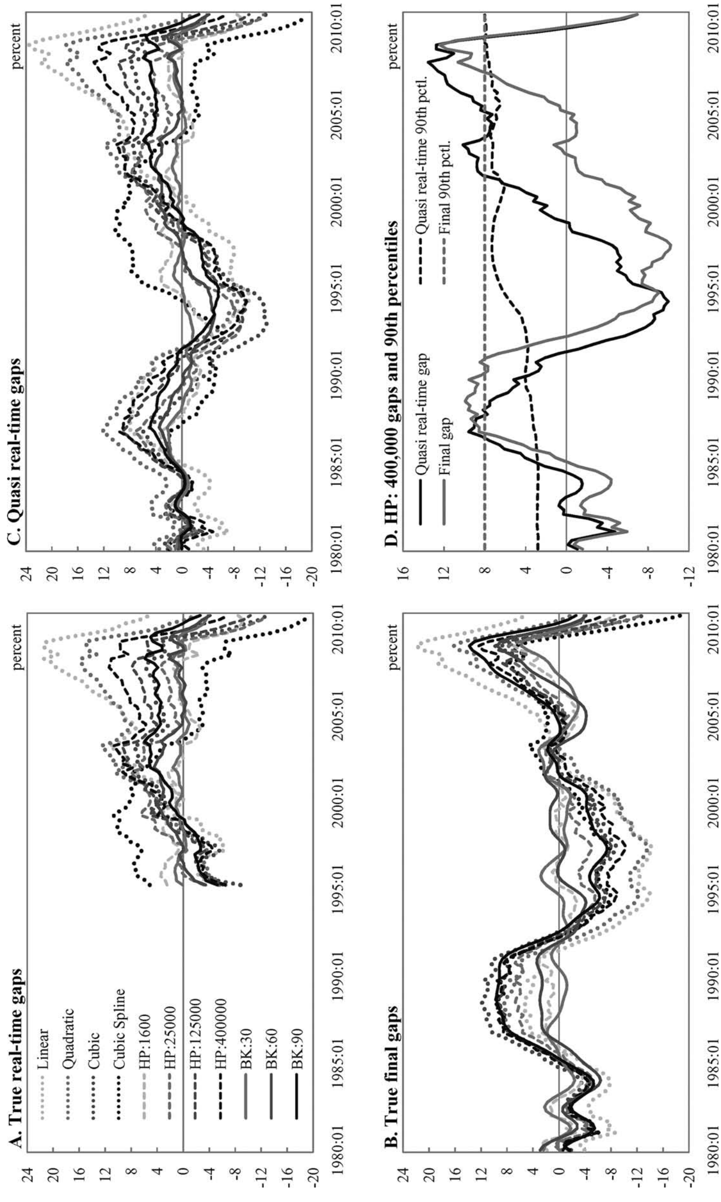
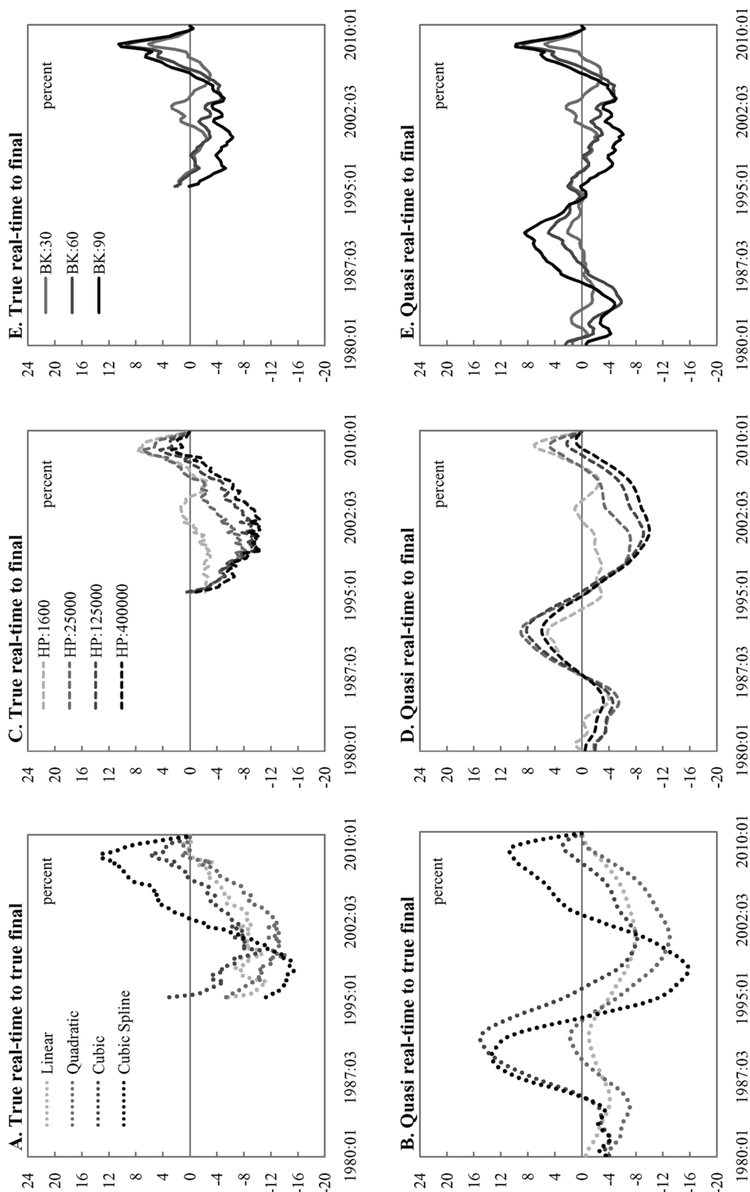


Figure 3. Credit-to-GDP Ratio Gap Revisions



was at some earlier point in time and therefore of the gap. Comparing the quasi real-time to final revisions with the true real-time to final revisions allows us to identify the portion of the overall revision that stems from *revisions* to the data and the portion that stems from the *availability* of data for later periods. Here, we find that the main source of revisions is not revisions to the underlying data but rather revisions caused by the unreliability of end-of-sample estimates of the trend. This is an important result for understanding why real-time estimates of credit-to-GDP ratio gaps are unreliable, and underscores the fact that different vintages of the credit-to-GDP ratio that do not revise much does not imply that the *gap* will not revise substantially. This result closely resembles Orphanides and van Norden's (2002) findings for output-gap estimates.

For the analysis that follows, the result that almost all of the revisions to the credit-to-GDP ratio gap stems from problems with end-of-sample trend estimation (and not from data revisions) means that we can focus on the quasi real-time to final revisions—for which we have a longer sample period—without worrying about missing a sizable source of revisions. We therefore now consider the quasi real-time to final revisions shown in the lower panels of figure 3.

4.2 Revision Magnitudes

Comparing the lower panels of figure 3 with the credit-to-GDP ratio gaps shown in figure 2 indicates that, with the exception of the linear trend model (which has extremely large gaps towards the end of the sample), the revision magnitudes are about the same order of magnitude as the gap estimates themselves. This is also evident in table 1, and the upper panels of tables 2 and 3, which report key summary statistics for the gap estimates and their revisions.

Table 1 reports the mean, standard deviation, and minimum and maximum values for the final and quasi real-time gaps implied by all of the filtering methods over the period 1980:Q1 to 2010:Q4. (Because the sample period for the data begins twenty-six years earlier, the means shown in the table need not equal zero.) The upper panel of table 2 reports similar statistics for the revisions across these gap estimates and for the persistence of these revisions.

The upper panel of table 3 presents results that assess the reliability of the quasi real-time gap estimates. The first two columns

Table 1. Credit-to-GDP Ratio Gap Summary Statistics

Method	Mean	Std. Dev.	Minimum	Maximum	Corr. w/Final
Linear					
Final	−1.08	9.09	−14.20	21.95	
Quasi Real Time	2.89	8.71	−10.32	23.51	0.97
Quadratic					
Final	−0.80	7.65	−13.60	16.21	
Quasi Real Time	4.99	6.86	−8.52	17.90	0.81
Cubic					
Final	0.46	6.27	−12.97	11.89	
Quasi Real Time	−0.15	6.15	−13.07	8.00	0.37
Cubic Spline					
Final	0.10	5.90	−18.74	9.46	
Quasi Real Time	0.15	5.95	−18.74	10.42	−0.11
HP: 1,600					
Final	−0.01	2.30	−8.47	6.98	
Quasi Real Time	−0.02	2.69	−9.33	3.46	0.41
HP: 25,000					
Final	−0.01	4.20	−12.55	7.81	
Quasi Real Time	0.92	4.74	−12.55	7.75	0.45
HP: 125,000					
Final	0.04	5.52	−10.35	9.89	
Quasi Real Time	1.92	5.77	−10.35	9.83	0.61
HP: 400,000					
Final	−0.10	6.35	−10.21	12.39	
Quasi Real Time	2.45	6.42	−9.99	13.53	0.73
BK: 6 to 30					
Final	0.04	1.87	−3.86	5.80	
Quasi Real Time	0.07	1.14	−3.86	2.27	0.56
BK: 6 to 60					
Final	0.05	3.22	−6.41	9.52	
Quasi Real Time	0.35	2.34	−5.06	4.17	0.41
BK: 6 to 90					
Final	1.03	5.93	−7.41	13.80	
Quasi Real Time	1.13	3.25	−5.69	5.97	0.68

Table 2. Revision Summary Statistics

Method	Mean	Std. Dev.	RMSE	Minimum	Maximum	AR
Entire Quasi Real-Time to Final Revision						
Linear	−3.97	2.38	4.63	−7.96	0.00	1.000
Quadratic	−5.79	4.49	7.32	−13.03	1.79	1.002
Cubic	0.61	6.97	6.97	−8.21	15.09	0.995
Cubic Spline	−0.05	8.84	8.80	−15.85	13.34	0.994
HP: 1,600	0.02	2.74	2.73	−4.02	7.05	0.981
HP: 25,000	−0.93	4.72	4.79	−7.21	9.02	0.994
HP: 125,000	−1.87	5.01	5.33	−9.14	8.19	0.997
HP: 400,000	−2.56	4.73	5.36	−10.02	5.92	0.997
BK: 6 to 30	−0.03	1.55	1.54	−2.94	5.48	0.932
BK: 6 to 60	−0.30	3.11	3.11	−5.85	8.76	0.970
BK: 6 to 90	−0.10	4.40	4.39	−6.12	9.82	0.985
Revisions within One Year						
Linear	−0.20	0.75	0.78	−1.60	0.99	0.989
Quadratic	−1.02	1.35	1.69	−3.04	1.76	0.992
Cubic	0.09	2.13	2.12	−2.40	4.90	0.997
Cubic Spline	−0.33	3.44	3.44	−6.33	8.86	1.022
HP: 1,600	−0.05	1.75	1.74	−2.34	6.43	1.034
HP: 25,000	−0.44	1.83	1.87	−3.01	4.44	1.016
HP: 125,000	−0.59	1.59	1.69	−2.63	2.73	1.000
HP: 400,000	−0.57	1.36	1.47	−2.74	2.08	0.992
BK: 6 to 30	−0.04	1.29	1.29	−1.69	5.16	0.884
BK: 6 to 60	−0.05	1.87	1.86	−2.56	7.70	0.882
BK: 6 to 90	−0.05	2.01	2.00	−2.79	8.41	0.881

quantify the earlier visual evidence on the magnitudes of the revisions—i.e., the "noise" present in the estimates—relative to the magnitude of the (final) measure itself—i.e., the "signal" from the estimates. The first column reports the ratio of the standard deviation of the revision to the standard deviation of the credit-to-GDP ratio gap, while the second column reports the ratio of the root mean squared error (RMSE) of the revision to the standard deviation of the credit-to-GDP ratio gap; the difference between these two measures is that the latter reflects any bias that is present in

Table 3. Summary Reliability Indicators

Method	Noise to Signal (1)	Noise to Signal (2)	Corr.	Opposite Sign	In Latter 90 Pctl. if in QRT 90 Pctl.	In QRT 90 Pctl. if in Latter 90 Pctl.
Entire Quasi Real-Time to Final Revision						
Linear	0.26	0.51	0.97	0.11	0.39	1.00
Quadratic	0.59	0.96	0.81	0.37	0.32	1.00
Cubic	1.11	1.11	0.37	0.39	0.21	0.30
Cubic Spline	1.50	1.49	−0.11	0.66	0.03	0.04
HP: 1,600	1.19	1.19	0.41	0.38	0.15	0.33
HP: 25,000	1.12	1.14	0.45	0.40	0.19	0.39
HP: 125,000	0.91	0.97	0.61	0.31	0.33	0.70
HP: 400,000	0.74	0.84	0.73	0.27	0.40	0.91
BK: 6 to 30	0.83	0.82	0.56	0.30	0.10	0.05
BK: 6 to 60	0.97	0.97	0.41	0.32	0.00	0.00
BK: 6 to 90	0.74	0.74	0.68	0.23	0.31	0.22
Revisions within One Year						
Linear	0.08	0.09	1.00	0.01	0.92	1.00
Quadratic	0.18	0.22	0.99	0.03	0.88	1.00
Cubic	0.34	0.34	0.98	0.03	0.38	1.00
Cubic Spline	0.58	0.58	0.87	0.17	0.05	0.67
HP: 1,600	0.76	0.76	0.60	0.21	0.12	0.45
HP: 25,000	0.44	0.45	0.96	0.08	0.43	0.95
HP: 125,000	0.29	0.31	0.99	0.03	0.46	0.96
HP: 400,000	0.21	0.23	1.00	0.03	0.62	1.00
BK: 6 to 30	0.69	0.69	0.52	0.27	0.40	0.24
BK: 6 to 60	0.58	0.58	0.73	0.25	0.55	0.41
BK: 6 to 90	0.34	0.34	0.85	0.12	0.75	0.48

the quasi real-time estimate. (The only notable differences between these two noise-to-signal ratios occur for the gaps implied by linear and quadratic detrending.)

Consistent with the comparison of figures 2 and 3, the noise-to-signal ratios of the gap estimates are high. The noise measures for the gap estimates in real time equal 75 percent to 150 percent of the size of the signal for every estimate except those implied by linear and quadratic detrending. (For these measures it is not wholly the case that the standard deviations of the revisions are smaller; much of this result is driven by the range of the gap estimates being so large.)

The remaining columns in the upper half of table 3 gauge the difference in signal across the quasi real-time and final gap measures. Specifically, the third column reports the correlations between the quasi real-time and final gap measures, and the fourth column reports the fraction of times that the gap estimates take on different signs. With the exception of the gap implied by the linear trend, the quadratic trend, and to some extent the HP-filtered trend with $\lambda = 400{,}000$, the correlations between the quasi real-time and final gaps are relatively low—on the order of 0.35 to 0.7 (this range excludes the cubic spline, for which the correlation is negative). The number of times that the quasi real-time and final gaps have opposite signs is also quite high—on the order of 25 to 40 percent of the time for most gap estimates. Gap estimates have the opposite sign relatively less frequently for the linear trend (about 10 percent of the time) but have the opposite sign relatively more frequently for the cubic-spline trend (about two-thirds of the time).

4.3 Revisions and Ex Post Policy Actions

In using the credit-to-GDP ratio gap to guide policy, an important question is whether the gap accurately signals in real time that policymakers should be requiring banks to accumulate capital buffers. To look at this question, we examine whether the quasi real-time and final estimates of the gap typically lie in the upper portions of their respective distributions in identical quarters.

Panel D of figure 2 shows how we answer this question for the HP-filtered trend with λ set to 400,000. The thick and thin solid black lines shown in panel D are, respectively, the quasi real-time and final gaps implied by this filtering method. The thick dashed black line gives what would have been considered the 90th percentile credit-to-GDP ratio gap in quasi real time. We calculate this series iteratively, in the same manner that we calculate the quasi real-time gaps; that is, we extend the sample period-by-period, calculate the gap series and its 90th percentile for that sample, and then combine each 90th percentile estimate into the single series shown in the figure. (Because the time series of gaps changes with each additional period added to the sample, this 90th percentile series also changes over time.) Finally, the thin dashed line in the figure is the 90th percentile of the final credit-to-GDP ratio gap; this is the estimate of the 90th percentile of the gap based on all information available up to the end of 2010.

If we take the 90th percentile of the credit-to-GDP ratio gap to be the level at which supervisors would deploy countercyclical capital buffers—which is consistent with the more frequent extreme of the ten- to twenty-year incidence described in the MVTF's consultative document—we can ask in which periods countercyclical capital buffers would have been in place in quasi real time.[6] For the HP filter with $\lambda = 400{,}000$ (panel D of figure 2), these would be the quarters in which the thick solid line (representing the gap) exceeds the thick dashed line (representing the 90th percentile). We can then ask how often for these instances that we also find that the final estimate of the gap exceeds the 90th percentile. The second-to-last column of table 3 gives this percentage, which in some cases is near zero and never exceeds 40 percent. The last column of table 3 then asks for what proportion of the time that the final gap is found to exceed the 90th percentile was the quasi real-time gap also found to be in the 90th percentile. Here we find higher numbers for almost all of the gap measures (except for those generated by the band-pass filter). This indicates that, for these detrending methods, reacting to levels of

[6]The MVTF's consultative document does not suggest using a percentile of the gap series as the threshold for deployment; rather, it indicates a 2 percent threshold (albeit adjusted to reflect the particular filter being used). We view a set percentile as a convenient way to define the threshold given the ranges of our different gap estimates. We consider this approach to be somewhat akin to the adjustment the MVTF suggested for different filters, although the threshold implied by the 90th percentile is generally several percentage points above the MVTF's 2 percent cut-off.

credit that appear to be excessive but later turn out not to be is a greater problem than missing—and thereby failing to react—when credit levels are excessive. With the band-pass filter, however, the opposite is the case.

It is evident from table 3 and panel D of figure 2 that there are occasions when quasi real-time estimates of the credit-to-GDP ratio gap are very high (specifically, above the quasi real-time estimate of the 90th percentile), even though the final estimates are not excessive by this definition. Figure 4 summarizes the timing of these occasions across all of the filtering methods that we consider. In the upper panel we assign a value of one to all periods in which the quasi real-time gap estimate has a value that puts it above the 90th percentile of the historical distribution of gaps. For each period, we then sum across the different filtering methods that put the quasi real-time gap in the 90th percentile. In the lower panel we do the same thing for the final gap. In the 2001 to 2003 period, many filtering methods indicate that countercyclical capital buffers should have been deployed based on quasi real-time estimates (the upper panel) but not based on full-sample estimates (the lower panel).[7] We consider in section 5 what implications these "false positives" might have had for lending. In addition, we would note that for these years the *true* real-time gap estimate would also have resulted in the deployment of countercyclical capital buffers.

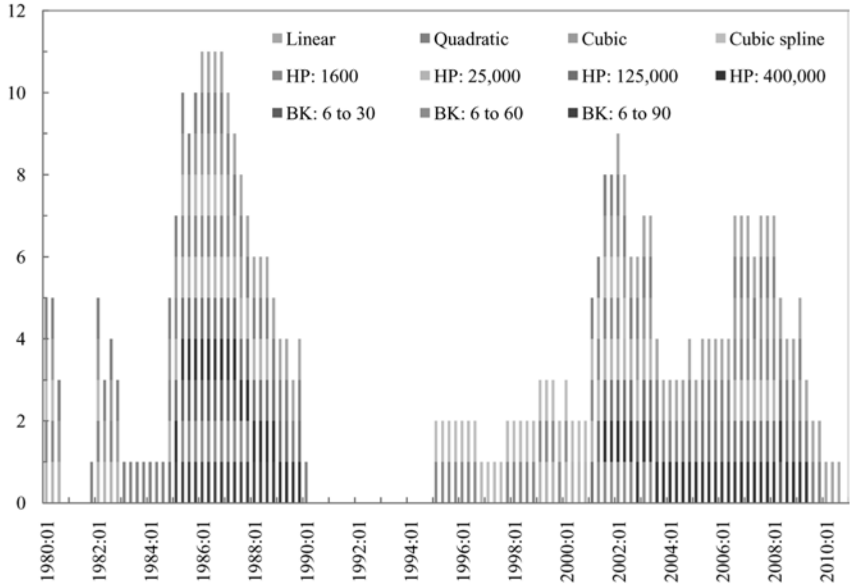
4.4 Within-Year Revisions

The MVTF's countercyclical capital buffer proposal gives banks one year to accumulate sufficient capital to meet additional regulatory requirements. This raises the concern that if estimates of the credit-to-GDP ratio gap revise substantially within a given year, it may weaken supervisors' ability to effectively deploy countercyclical capital buffers. The lower panels of tables 2 and 3 report statistics that

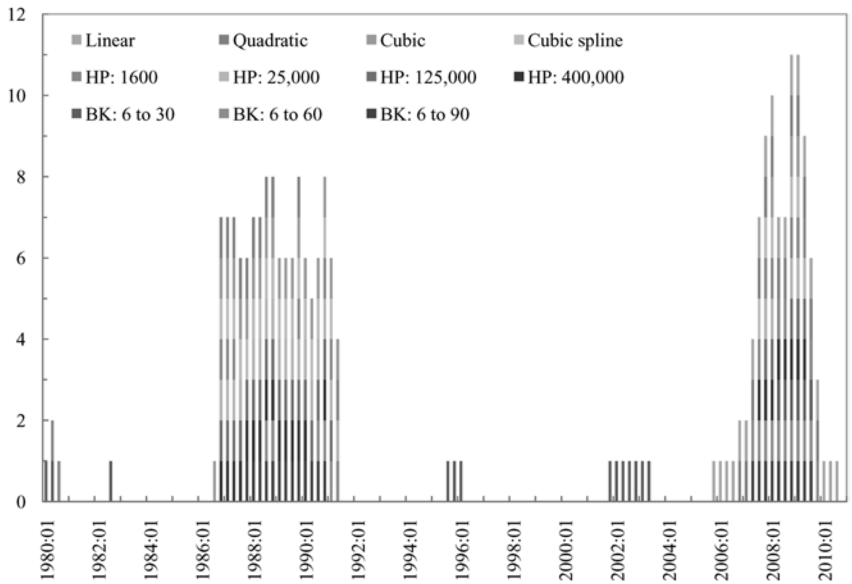
[7]One problem that has been noted with regard to using elevated credit-to-GDP ratio gaps to signal the deployment of countercyclical capital buffers is that this ratio can rise in economic downturns purely because GDP typically declines relatively faster than credit. (See the MVTF's consultative document and Repullo and Saurina 2011.) Replicating our exercise using the credit-to-potential-GDP ratio gap (where we use the Congressional Budget Office's February 2011 estimate of potential GDP) does not materially alter our results, however.

Figure 4. Quarters in which the Credit-to-GDP Ratio Gap Is in the 90th Percentile

A. Quasi real-time



B. Final



address this issue. The first two columns of the lower panel of table 3 indicate that revisions to the quasi real-time gap over the first year after the estimate is made are modest for most filtering methods. The exceptions, however, are the HP-filtered trend with a λ value of 1,600 and the band-pass filters that allow periodicities of six to thirty periods and six to sixty periods, for which two-thirds of the ultimate revision occurs within the first year. (The results of the third and fourth columns are similar.)

The last two columns of this panel consider how the decision to deploy countercyclical capital buffers could change given an extra year of data with which to estimate the gap. As before, we continue to find that false positives are a larger problem for the deterministic detrending methods and the HP filter, while missed responses appear to be a greater concern for the band-pass filter.

5. Revisions and Different Filtering Methods

The relationship between the choice of filtering method and the size and timing of gap-estimate revisions can be summarized as follows. First, the credit-to-GDP ratio gaps obtained from the linear and (to a lesser extent) quadratic detrending methods yield the smallest revisions. Second, deterministic detrending methods with higher-order polynomial trends imply larger revisions, while the HP-filter and frequency detrending methods that allow for longer credit cycles imply smaller revisions. Finally, gap estimates implied by frequency detrending methods exhibit a larger portion of their revisions quite quickly—that is, within the course of a year—relative to other methods.

The finding that credit-to-GDP ratio gaps implied by linear and quadratic detrending methods have the smallest revisions reflects the problems that other detrending methods have at the endpoints of a series. These problems arise because both unobserved-components and band-pass filtering methods estimate the trend and cycle of a series using past and current values of a series, as well as whatever future observations are available. Toward the end of the sample period, fewer future observations are available, which can result in large revisions when data do become available.

The Baxter-King filter, which is implemented as a symmetric two-sided moving average of the actual data, uses an estimated AR process to "pad" the sample period with backcasts and forecasts of the series to be filtered. As additional quarters of actual data become available, the future-period observations and forecasts used in the filter also change, potentially resulting in large revisions to the gap and trend estimates. Of course, the degree to which these estimates of the trend and cycle change depends on the weights in the moving average assigned to the quarters for which new data and revised forecasts become available. The HP filter, which can also be expressed as a two-sided moving average of the actual data, does not pad the sample period with backcasts and forecasts but instead applies different weights at different points in the sample period. That is, if we are calculating the cyclical component of a series for the first period of the sample, the coefficients in the moving average applied to the actual data are different from the coefficients applied for the second period in the sample, the third period in the sample, and so on. The amount by which the moving-average coefficients change from one observation to the next depends on where in the sample the observation lies: At the beginning and end of a sample, the moving-average weights differ greatly from one observation to the next, while in the middle of a large sample the moving-average weights change little. As Baxter and King (1999) discuss, the moving-average weights for an HP filter with $\lambda = 1{,}600$ only settle down after about three years.[8] This yields significant instability in our real-time estimates of the credit-to-GDP ratio's trend and associated gap measure.

Deterministic detrending methods model the trend as the fitted value of an estimated polynomial function of time, with the cyclical component defined as the residual between the series and the trend. Here, additional quarters of data result in revisions to the trend and the cycle through their effects on the estimated coefficients of the polynomial. In principle, additional quarters of data should alter the trend and cycle only very modestly in large samples because additional observations should yield only slight changes in parameter

[8]Baxter and King also document that in the early and late parts of the sample the HP filter is not a good approximation to the high-pass filter (see subsection 2.3 above). Three years into the sample and three years from the sample's end, however, the HP filter is a better approximation.

values. The results in tables 2 and 3 indicate that this is in fact the case for the credit-to-GDP ratio when we use linear or quadratic detrending. However, when we model the trend as a cubic function of time, the real-time reliability of the resulting gap measure is among the poorest of all the filtering methods we consider. This in turn reflects overfitting at the endpoint: As can be seen from the plot of the credit-to-GDP ratio in figure 1, there is a large run-up and subsequent decline in the series over the early 1980s to early 1990s period. A higher-order polynomial initially attempts to fit this bulge with a small increase in the trend at the end of the sample. As more data become available and the run-up starts to reverse itself, the cubic polynomial calls for a flatter trend, which in turn implies large revisions to the gap. The reason this does not happen for the linear and quadratic trends is that they never attempt to fit the bulge in the credit-to-GDP ratio; hence their parameter estimates change relatively little around this episode and smaller revisions obtain. This underscores the sensitivity of even relatively low-order polynomial detrending procedures in real time when the actual series exhibits persistent but ultimately transitory movements.

The cubic spline has the largest revisions of all the methods we consider. This reflects both the problems faced by cubic polynomial detrending and the fact that the estimation intervals for the segments of the spline can be quite small despite a large available time series of data. Specifically, our time series of twenty-seven years at the start of our real-time analysis and fifty-seven years over the complete sample translate into spline segments that initially span nine years of data and eventually span nineteen years. Consequently, additional quarters of data can have significant effects on the trend and result in large revisions to the estimated gap.

We would note that although the linear and quadratic trends exhibit the smallest revisions, they are not necessarily the best techniques to use for estimating the credit-to-GDP ratio gap. An augmented Dickey-Fuller test of the credit-to-GDP ratio over the full sample indicates that the series has a unit root and thus has a stochastic trend. This implies that the HP filter and the band-pass filter that are able to remove a unit root are more appropriate to apply to the credit-to-GDP ratio since deterministic detrending methods will generate spurious cycles. Although our focus is on revisions to the various gap estimates, any practical attempt to use a credit-to-GDP ratio gap to guide countercyclical capital policy would require some consideration of issues such as the deterministic or stochastic nature of the trend and the most appropriate filter to use.

Another feature of the estimated revisions implied by the various detrending procedures is that the HP-filter and frequency detrending methods that allow for longer credit cycles imply smaller revisions to the implied gap measures. For the HP filter, this arises because longer assumed credit cycles imply a smoother path of the trend (this is associated with a larger penalty on changing the slope of the trend in the HP-filter optimization problem—see subsection 2.2). This means that additional observations have a smaller effect on the estimated cyclical component with correspondingly smaller revisions to the gap measure. The intuition for the Baxter-King filter is similar: Allowing for longer credit cycles implies that a smaller range of low-frequency fluctuations are extracted in constructing the cycle; additional observations therefore have smaller effects on the low-frequency component of the series, in turn implying smaller revisions to the cyclical component.

Finally, the result that a larger share of revisions occur within a year if we use filtering methods with shorter credit cycles (e.g., the HP-filtered trend with a λ of 1,600 and the band-pass filters that pass through periodicities in a range of six to thirty quarters) reflects the smaller amounts of data that are needed to change the trend estimates for these filtering methods. Similarly, the different ways that the HP and band-pass filters deal with endpoint problems determines the rapidity with which the bulk of the revisions occur. As noted earlier, it takes about three years before the moving-average weights associated with the HP filter settle down; hence, a reasonable fraction of revisions to the gap occur more than one year after the reference quarter. Revisions to the Baxter-King gaps occur more quickly because the largest moving-average weights are on observations that are just a couple of quarters before and after the current observation. Consequently, the largest revisions for the Baxter-King gaps occur within a year.

6. Real Implications of Countercyclical Capital Buffers

The results of section 3 underscored a key practical difficulty associated with countercyclical capital buffers—specifically, the tendency for credit-to-GDP ratio gaps to yield false positives in terms of indicating excessively high levels of credit in real time. We now consider the potential real economic costs of deploying countercyclical capital buffers based on unreliable real-time estimates of the credit-to-GDP ratio. We focus on the reduction in lending that would have obtained were countercyclical capital buffers to have been deployed in 2001:Q4 and 2003:Q2, which are the dates for which a number of detrending methods yielded false positives (see subsection 3.3). The effect of countercyclical capital requirements on lending in these quarters depends on a number of considerations, including the extent to which countercyclical capital requirements would have been binding in these quarters, the capital shortfalls that countercyclical buffers would have implied, and the effect of increased capital requirements on the level of lending and interest rates.

While our paper focuses on the costs of potential false positives, this should not be read as necessarily implying that countercyclical capital buffers carry no potential benefits. The MVTF's consultative document suggested one such benefit, which is that if a credit boom associated with an increase in systemic risk is under way, the deployment of countercyclical capital buffers should leave banks better positioned to absorb losses—which should in turn reduce the risk that regulatory capital requirements will constrain lending when credit market conditions deteriorate. The document also contends that the build-up of capital buffers carries a possible side benefit in that it might retard the expansion of excess credit. Although the consultative document makes no attempt to demonstrate the existence of these benefits (either in the context of a structural model or through past natural experiments), we are broadly sympathetic to these claims on a priori grounds. We focus on the *costs of false positives* because this addresses the considerations that follow most naturally from our evidence for the *incidence of false positives* in section 3. Moreover, descriptions of the purported potential benefits of countercyclical capital buffers are available elsewhere; see Caruana (2010) for a discussion in the context of the run-up to the 2007–09 crisis.

6.1 Extent that Countercyclical Capital Requirements Bind

We examine the extent to which countercyclical capital buffers deployed in 2001:Q4 and 2003:Q2 would have been binding by considering the distribution of risk-based capital (RBC) ratios across U.S. banks for these two quarters. The left column of figure 5 shows the distribution of unweighted RBC ratios in these quarters by institution and the right column shows the distribution of weighted RBC ratios by asset volume. As can be seen from the left column of the figure, in 2001:Q4 and 2003:Q2 nearly all banks met the Federal Deposit Insurance Corporation's (FDIC's) 10 percent total RBC ratio criterion for being well capitalized.[9] The right column, however, shows much more bunching in RBC ratios when the ratios are weighted by a bank's total assets, with most of the distribution lying just above the 10 percent cut-off for being well capitalized.

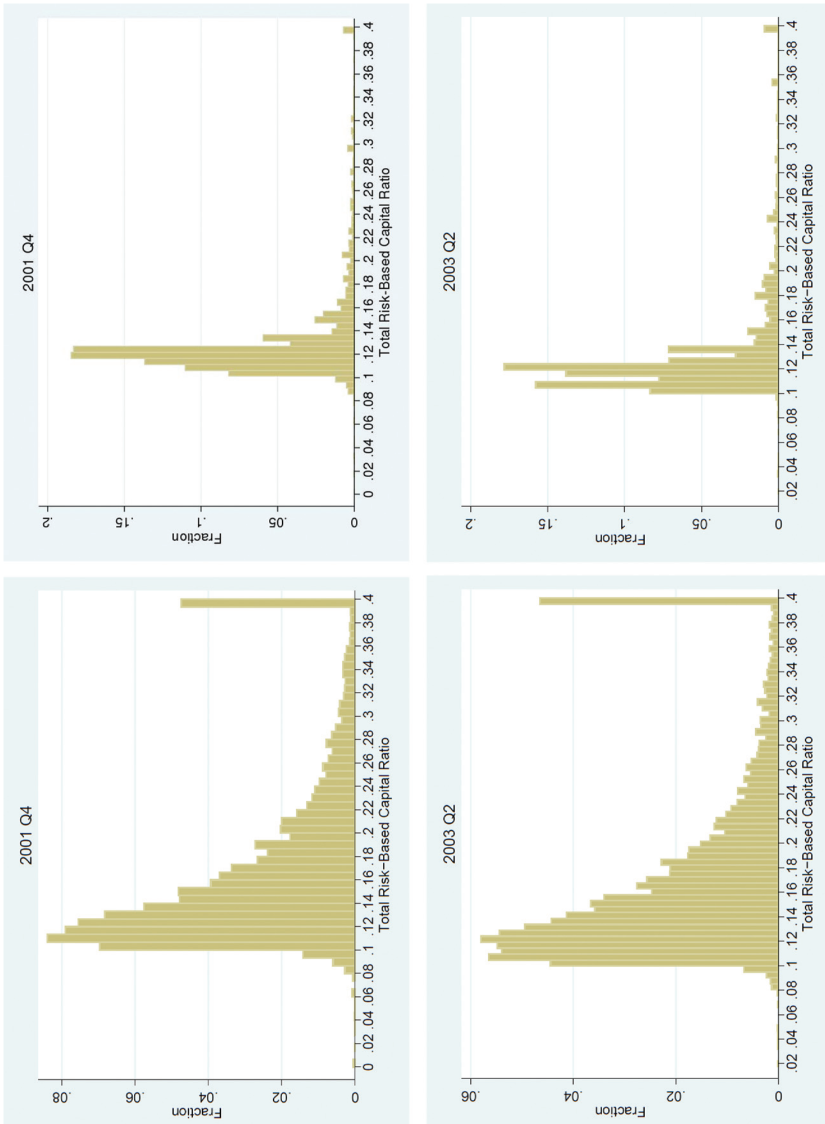
Two observations follow from figure 5. First, the changes in capital ratio requirements that would have been implied by credit-to-GDP ratio measures in 2001:Q4 and 2003:Q2 would have been binding given that banks strive to keep their capital ratios in the well-capitalized category.[10] Second, large banks, which account for the lion's share of assets and capital in the banking sector, have capital ratios that are only slightly above regulatory requirements.[11] In sum, these observations imply that changes in capital ratio requirements based on credit-to-GDP ratio measures in 2001:Q4 and

[9]The FDIC defines a bank to be well capitalized if its total RBC ratio is at least 10 percent, its tier 1 RBC ratio is at least 6 percent, and its leverage ratio is at least 5 percent; it is adequately capitalized if it is not well capitalized and its total RBC ratio is at least 8 percent, its tier 1 RBC is at least 4 percent, and its leverage ratio is at least 4 percent; and it is undercapitalized if it fails to meet the adequately capitalized criteria. That said, a leverage ratio of at least 3 percent suffices for being defined as adequately capitalized if a bank's CAMELS rating is 1 and the bank is not experiencing significant growth. A bank's assessed contribution to the deposit insurance fund is based on its capitalization category.

[10]This finding contradicts Hanson, Kashyap, and Stein (2011), who argue that during recessions regulatory capital requirements are not binding. Rather, they contend that the minimum capital ratio that binds in recessions is the one imposed by markets, which become willing only to fund very strongly capitalized banks. Therefore, they assert that regulatory capital requirements during good times need to exceed the market's capital requirements in recessions to achieve the desired countercyclical effect, and they suggest a range of 12 to 15 percent—well above the maximum capital requirements in the MVTF's consultative document.

[11]This second fact is consistent with Hanson, Kashyap, and Stein (2011), who argue that competition puts pressure on banks to reduce capital ratios. Although smaller banks tend to have higher capital ratios, they also appear to utilize different lending technologies, most notably relationship lending (see Berger and Black 2011 and Petersen and Rajan 1994).

Figure 5. Distribution of Unweighted and Weighted Capital Ratios



2003:Q2 would have affected a *significant share* of the economy's bank capital, even if the changes would have affected only a *small number* of institutions.

6.2 Capital Shortfalls Implied by Countercyclical Capital Requirements

The data underlying figure 5 allow us to compute how much capital the banking sector would have needed to raise in 2001:Q4 and 2003:Q2 to meet the capital requirements imposed by countercyclical capital buffers. This calculation is done by determining for each quarter which banks would have had a shortfall of capital relative to the new countercyclical capital buffer requirements, calculating the dollar value of the shortfall for each bank, and summing these individual bank shortfalls together to obtain the aggregate shortfall for the entire banking sector. The upper part of table 4 shows the outcome of these calculations for a range of countercyclical capital buffer add-ons; specifically, for a .5-percentage-point, 1-percentage-point, 1.5-percentage-point, and 2-percentage-point add-on. (We do not use the maximum possible add-on of 2.5 percentage points so as not to obtain extreme results with regard to the potential costs we calculate.)

It is well documented that banks operate with higher capital ratios than are required by regulation, with many factors—including the capital ratio demanded by credit markets, managers' preferences for risk management, and the regulatory and supervisory environment—accounting for the difference between the actual and required capital ratios. (See Alfon, Argimon, and Bascuñana-Ambros 2004.) Thus, we also considered the amount of the capital shortfall in the banking system were banks to hold a precautionary buffer in excess of the higher regulatory minimum. Specifically, we considered the capital shortfall were supervisors to deploy a 2-percentage-point countercyclical capital buffer under the assumption that banks want to hold an additional 1/2 percent or 1 percent of capital. (These estimates are also shown in the top panel of table 4.)

Our calculations suggest the following ranges of estimates for the capital shortfalls that would be implied by countercyclical capital buffers. In 2001:Q4, the total capital shortfall of U.S. banks would have been between $2.1 billion for a capital ratio requirement of 10.5

Table 4. Counterfactual Capital Shortfall and Reduction in Lending

Required Total Risk-Based Capital Ratio	Capital Shortfall (in billions) 2001:Q4	2003:Q2
10.5%	2.1	1.1
11.0%	5.5	6.3
11.5%	12.9	15.7
12.0%	24.4	28.4
12.0% + 0.5% Precautionary Buffer	39.9	46.4
12.0% + 1.0% Precautionary Buffer	58.8	67.3

	Reduction in Lending (in billions)			
	2001:Q4		2003:Q4	
Required Total Risk-Based Capital Ratio	Lower Bound	Upper Bound	Lower Bound	Upper Bound
10.5%	3.9	26.7	2.0	30.6
11.0%	10.2	52.6	11.7	61.0
11.5%	24.0	90.9	29.2	106.7
12.0%	45.4	141.1	52.8	162.1

Notes: The capital shortfall is defined as the total amount of capital needed by banks that have capital ratios below the requirements holding assets constant. The lower bound is constructed using the estimates for bank-specific target ratio of $1.86 of lending for $1 of capital as reported in Berrospide and Edge (2010). The upper bound is constructed using the estimates for regulatory capital shortfalls of $3.16 of lending for $1 of capital as reported in Hancock and Wilcox (1993) for each $1 of capital shortfall plus a 1 percent precautionary bank-specific target buffer with $1.86 of lending for $1 of capital.
Source: Call Reports.

percent and $58.8 billion for a capital ratio requirement of 12 percent together with a 1 percent precautionary buffer. For 2003:Q2, our estimates are $1.1 billion for a capital ratio requirement of 10.5 percent and $67.3 billion for a capital ratio requirement of 12 percent together with a 1 percent precautionary buffer.

6.3 Impact of Capital Shortfalls on Lending and Interest Rates

We first address the question of how the capital shortfalls reported in the upper panel of table 4 would have affected the dollar value of bank lending. There is a wide range of theoretically possible values for the scale of the increase; these range from zero (if banks can costlessly raise equity) to an amount equal to the leverage rate (if banks target regulatory requirements and actively manage their assets). Adopting this latter view, which was quite prominent in the early years of the recent crisis (see Hatzius 2007, 2008), would lead to very substantial reductions in assets and loan volumes from a capital shortfall (specifically, we would obtain dollar declines in assets on the order of ten times the dollar value of the capital shortfall). We do not adopt this view because such large effects are at odds with what estimated models of bank lending indicate. For example, Hancock and Wilcox (1993, 1994) find that a $1 capital shortfall decreases lending by $1.50; more recently, Berrospide and Edge (2010) estimate the effect of bank capital on lending using bank holding company data and confirm the modest effects of Hancock and Wilcox. In their data, Berrospide and Edge find that a $1 shortfall in bank capital results in a $1.86 reduction in loans.[12]

The estimates of Berrospide and Edge (2010) reported above focus on bank capital shortfalls in general and not on capital shortfalls following a change in regulatory capital requirements. Hancock and Wilcox (1993) do distinguish capital shortfalls relative to bank-specific targets and relative to regulatory requirements, which arose following the implementation of the Basel Accord. The latter have significantly larger effects on lending, with a $1 regulatory capital shortfall yielding a $3.16 reduction in total lending.

We construct a lower and upper bound for the implications of capital shortfalls implied by countercyclical capital buffers being deployed in 2001:Q4 and 2003:Q2 on lending. The lower bound is the systemwide capital shortfall, excluding any precautionary buffers, multiplied by the $1.86 estimate of Berrospide and Edge for the effect

[12]See also Bernanke and Lown (1991), who estimate a 2- to 2.5-percentage-point increase in loan growth for a 1-percentage-point increase in capital ratios. This estimate is also much less than what would be implied by banks targeting their regulatory requirements and actively managing their assets.

of capital shortfalls on loan volumes.[13] The upper bound assumes that banks also build up a precautionary buffer of 1 percentage point in excess of the new, higher regulatory minimum. Here we apply different estimates of the effects of capital shortfalls depending on whether the shortfall is a regulatory shortfall or is relative to a bank's own desired capital target: For regulatory shortfalls we apply the Hancock and Wilcox estimate of a $3.16 reduction in loan growth for a $1 capital shortfall, while for the shortfall relative to the bank's own desired capital ratio we apply the Berrospide and Edge estimate of $1.86.[14] The results for all of the capital ratio changes we consider are shown in table 4. Even without making any extreme assumptions, the effects of countercyclical capital buffers on lending can be substantial. For example, for an additional capital requirement of 2 percentage points, the reduction in lending following 2001:Q4 could have been as much as $141 billion.

Our calculations so far have focused on loan volumes. However, higher capital requirements can also affect banks' funding costs and translate into higher spreads for borrowers. That said, there appears to be little empirical evidence to support this possibility. For example, Meisenzahl (2011) finds no significant effect of funding costs or capital ratios on business loan interest rates in small business loan data. Specifically, for banks with more than $50 billion in assets in the 2003 Survey of Small Business Finances survey sample, Meisenzahl finds that a 10-percentage-point increase in the capital ratio increases the business loan interest rate only 23 basis points (this increase is insignificant). Similarly, Hanson, Kashyap, and Stein (2011) report a modest loan interest rate increase of at most 35 basis points for a 10-percentage-point increase in the capital ratio.[15] In sum, additional countercyclical capital buffers of 1 or 2 percentage points appear unlikely to increase loan interest rates by much.

---

[13]The $3.9 billion *lower-bound* lending decline shown in table 4 for a .5-percentage-point capital add-on is the product of the capital shortfall of $2.1 billion implied by regulation (in the top panel of the table) and $1.86.

[14]The $26.7 billion *upper-bound* lending decrease shown in table 4 for a .5-percentage-point capital add-on sums the product of the $2.1 billion capital shortfall implied by regulation and $3.16 with the product of the capital shortfall implied by banks' desired capital ratio (equal to $12.9 billion less $2.1 billion) and $1.86.

[15]In a companion paper, Kashyap, Stein, and Hanson (2010) report estimates in a range of 25 to 45 basis points for a 10-percentage-point increase in the capital ratio.

6.4 Some Context for the Decline in Lending

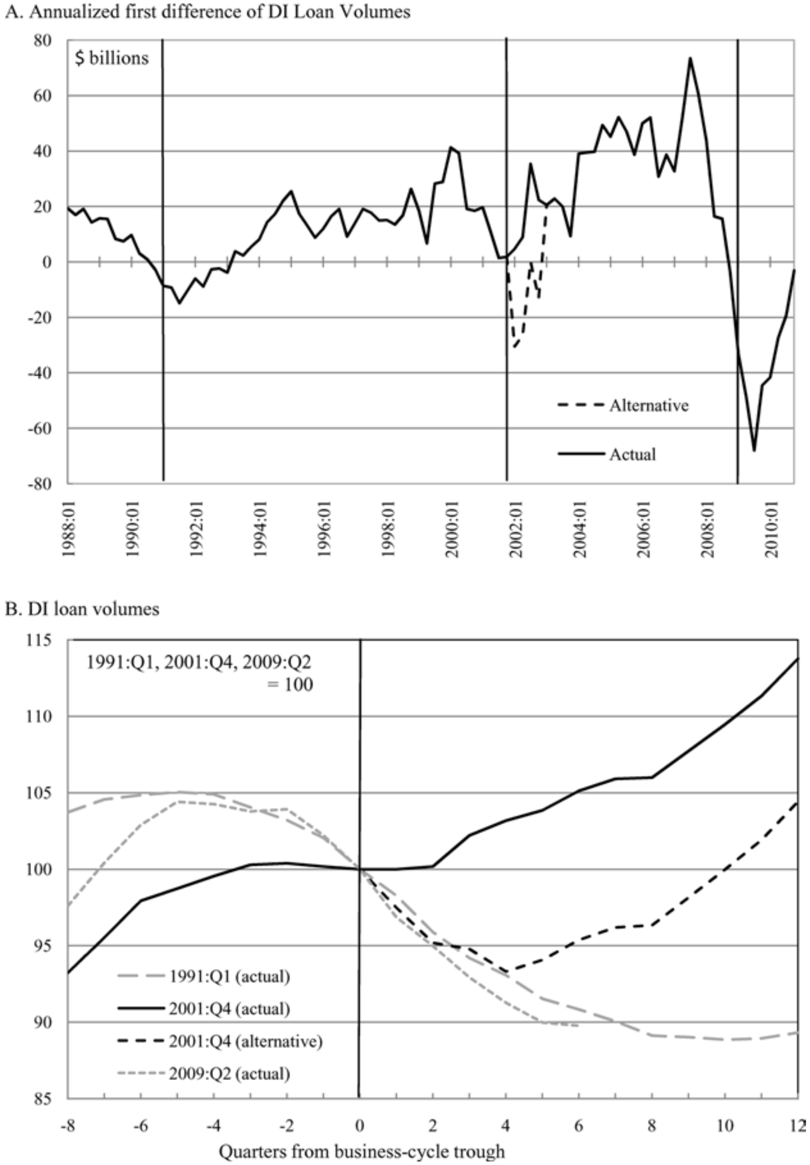
We now put the reductions in loan volumes reported in table 4 into context. The range of $3.9 billion to $141.1 billion for 2001:Q4 can be compared for the same time period with new mortgage originations to consumers of about $700 billion and new auto loan originations of about $100 billion, as well as with commercial and industrial (C&I) loan originations of about $88 billion for the first week of November 2001 ($1,114 billion for the quarter). Similarly, the range of $2 billion to $162.1 billion for 2003:Q2 can be compared for the same time period with new mortgage originations of almost $1,000 billion and new auto loan originations of about $75 billion, as well as with C&I loan originations of about $62 billion for the first week of May 2003 ($806 billion for the quarter).[16] If we assume that all of the reductions in loan volumes in table 4 occur in these bank-dependent loan categories (as research by Hancock and Wilcox 1993 and Peek and Rosengren 1995 suggests), originations could have been reduced by 7.6 percent and 8.9 percent of total loans in 2001:Q4 and 2003:Q2, respectively.

Figure 6 provides a different way of putting the decline in lending reported in table 4 in context. Here we focus on 2001:Q4, which is also the quarter that the National Bureau of Economic Research (NBER) dates as the trough of the 2001 recession. The solid line in panel A of figure 6 shows the *unannualized* first difference of the volume of depository institution (DI) loans over the period 1988 to 2010, as reported in the Flow of Funds Accounts. The dashed line subtracts $35.3 billion from the flow of DI loan volumes in each quarter of 2002. This amount equals one-quarter of the $141.1 billion reduction in lending that would have been implied by the deployment of a 2-percentage-point countercyclical buffer in 2001:Q4. (We spread the $141.1 billion amount across four quarters because the proposal in the MVTF's consultative document gives banks a year to increase their capital ratios.)

The solid black line in panel B shows the path of real (GDP-price deflated) loan volumes for DIs around 2001:Q4, rescaled to 100 in

[16]See the results of the Federal Reserve Bank of New York Consumer Credit Panel (FRBNY 2010), and the Survey of Terms of Business Lending, www.federalreserve.gov/releases/e2/.

Figure 6. Depository Institutions' Loan Volumes



that quarter. The dashed black line shows the (deflated and rescaled) path of loan volumes around 2001:Q4, assuming the deployment of a 2-percentage-point countercyclical capital buffer. As can be seen, the deployment of the buffer would have implied that over the first year following the 2001 recession, real DI loan volumes would have contracted, rather than remaining flat as they did in history.

To see whether such a contraction would have been significant, we overlay on panel B the paths of real DI loan volumes around the NBER troughs of the 1990–91 and 2007–09 recessions (which fell in 1991:Q1 and 2009:Q2, respectively). To make these lines comparable, we rescale them so that their values at the trough of each recession are 100. As is evident from the figure, had a 2-percentage-point capital add-on been put into effect in 2001:Q4 and had banks also chosen to hold an additional 1-percentage-point buffer, loan volumes in 2002 would have declined in a manner similar to how they contracted during the credit crunches that followed the 1990–91 and 2007–09 recessions.

To be sure, the MVTF's consultative document does not advocate a mechanistic rule in which the buffer is automatically deployed once the indicator variable exceeds some threshold. Rather, the document emphasizes the need for applying "judgment in the setting of the buffer . . . using the best information available to gauge the build up of systemic risk" (p. 7). Thus, it is unlikely that a buffer would be deployed without policymakers taking other economic conditions into account. Although the use of judgment could reduce the possibility of policymakers imposing countercyclical capital buffers when the economy is in recession, it certainly does not rule it out. Were systemic risks perceived—albeit by a false positive—to have been intensifying in the financial system in late 2001, policymakers could have still deemed it prudent to deploy buffers to ensure that banks would have enough capital to absorb losses once credit conditions deteriorated, even in the face of a downturn.[17]

[17]The MVTF's consultative document does not address how the effectiveness of a countercyclical capital buffer policy would be affected by allowing for a large degree of discretion in its implementation (for example, due to regulatory forbearance). While a triggering rule that also depends on macroeconomic conditions might be seen as a way to solve this, in practice such a rule would probably still permit a large degree of discretion given the well-known problems associated with determining the state of the economy in real time.

7. Conclusions

This paper has assessed the potential cost of the MVTF's proposal to use the credit-to-GDP ratio gap as a reference point for countercyclical capital buffers. Because these gap measures are very unreliable in real time, they provide a poor foundation for policymaking. Specifically, real-time measures of the gap can yield false positives by signaling excessively high levels of credit that later—based on longer time series of data—do not appear to be so extreme. When these measures are used to determine whether countercyclical capital buffers should be deployed, these false positives can in turn result in capital shortfalls in the banking sector and unnecessary lending restraint. We investigate a few instances in which the credit-to-GDP ratio gap does in fact yield a false positive, and find that in these episodes the impact on loan volumes can be highly significant.

To be clear, our paper is not a comprehensive evaluation of all potential costs, benefits, unintended consequences, and regulatory arbitrage possibilities associated with the Basel Committee on Banking Supervision's countercyclical capital buffer proposal. In particular, we do not attempt to quantify or provide an analytic demonstration of any of the benefits posited by the MVTF's consultative document; rather, our focus is solely on the potential risks of using an unreliable real-time indicator to guide countercyclical capital deployment decisions. Of course, basing decisions on potentially unreliable real-time indicators is a practical complication faced by any type of stabilization policy. But this does not mean that the costs of unreliable real-time indicators should be accepted as an inevitable consequence of policymaking. Rather, what is crucial is to design decisionmaking processes that are robust to the inaccuracies inherent in most real-time indicators and that therefore minimize the attendant costs of using these indicators to inform policy actions. The results of this paper indicate that tying the deployment of countercyclical capital buffers to the credit-to-GDP ratio gap does not meet these criteria.

References

Alessi, L., and C. Detken. 2009. "Real-Time Early Warning Indicators for Costly Asset Price Boom/Bust Cycles." ECB Working Paper No. 1039.

Alfon, I., I. Argimon, and P. Bascuñana-Ambros. 2004. "What Determines How Much Capital Is Held by UK Banks and Building Societies?" FSA Occasional Paper No. 22.

Basel Committee on Banking Supervision. 2010. "Countercyclical Capital Buffer Proposal." Consultative Document, Bank for International Settlements.

Baxter, M., and R. King. 1999. "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series." *Review of Economics and Statistics* 81 (4): 575–93.

Berger, A., and L. Black. 2011. "Bank Size, Lending Technologies, and Small Business Finance." *Journal of Banking and Finance* 35 (3): 724–35.

Bernanke, B., and C. Lown. 1991. "The Credit Crunch." *Brookings Papers on Economic Activity* 22 (2): 205–47.

Berrospide, J., and R. Edge. 2010. "The Effects of Bank Capital on Lending: What Do We Know, and What Does It Mean?" *International Journal of Central Banking* 6 (4): 5–54.

Borgy, V., L. Clerc, and J.-P. Renne. 2009. "Asset-Price Boom-Bust Cycles and Credit: What Is the Scope of Macro-Prudential Regulation?" Banque de France Working Paper No. 263.

Canova, F. 1998. "Detrending and Business Cycle Facts." *Journal of Monetary Economics* 41 (3): 475–512.

Caruana, J. 2010. "Macroprudential Policy: Could It Have Been Different This Time?" Speech given at the People's Bank of China Seminar on Macroprudential Policy, Shanghai, October 18.

Federal Reserve Bank of New York. 2010. "Quarterly Report on Household Debt and Credit: August 2010." Research and Statistics Group, Federal Reserve Bank of New York.

Hancock, D., and J. Wilcox. 1993. "Has There Been a 'Capital Crunch' in Banking?" *Journal of Housing Economics* 3 (1): 31–50.

———. 1994. "Bank Capital and Credit Crunch: The Roles of Risk-Weighted and Unweighted Capital Regulations." *Journal of the American Real Estate and Urban Economics Association* 22 (1): 59–94.

Hanson, S., A. Kashyap, and J. Stein. 2011. "A Macroprudential Approach to Financial Regulation." *Journal of Economic Perspectives* 25 (1): 3–28.

Hatzius, J. 2007. "Leveraged Losses: Why Mortgage Default Matters." *Goldman Sachs U.S. Daily*, November 15.

———. 2008. "Beyond Leveraged Losses: The Balance Sheet Effects of the Home Price Downturn." *Brookings Papers on Economic Activity* 39 (2): 195–227.

Hodrick, R., and E. Prescott. 1980. "Post-War U.S. Business Cycles: An Empirical Investigation." Manuscript, Carnegie-Mellon University.

Kashyap, A., J. Stein, and S. Hanson. 2010. "An Analysis of the Impact of Substantially Heightened Capital Requirements on Large Financial Institutions." Manuscript, Harvard University.

King, R., and S. Rebelo. 1993. "Low Frequency Filtering and Real Business Cycles." *Journal of Economic Dynamics and Control* 17 (1–2): 207–31.

Meisenzahl, R. 2011. "Verifying the State of Financing Constraints: Evidence from U.S. Business Credit Contracts." Finance and Economics Discussion Series Paper No. 2011-04, Board of Governors of the Federal Reserve System.

Peek, J., and E. Rosengren. 1995. "Bank Regulation and the Credit Crunch." *Journal of Banking and Finance* 19 (3–4): 679–92.

Petersen, M., and R. Rajan. 1994. "The Benefits of Lending Relationships: Evidence from Small Business Data." *Journal of Finance* 49 (1): 3–37.

Orphanides, A., and S. van Norden. 2002. "The Unreliability of Output-Gap Estimates in Real-Time." *Review of Economics and Statistics* 84 (4): 569–83.

Repullo, R., and J. Saurina. 2011. "The Countercyclical Capital Buffer of Basel III: A Critical Assessment." Manuscript, CEMFI.

Staiger, D., J. H. Stock, and M. W. Watson. 1997. "How Precise Are Estimates of the Natural Rate of Unemployment?" In *Reducing Inflation: Motivation and Strategy*, ed. C. Romer and D. Romer. Chicago: University of Chicago Press.

———. 2001. "Prices, Wages, and the U.S. NAIRU in the 1990s." In *The Roaring Nineties*, ed. A. Kruger and R. Solow. New York: Russell Sage Foundation.

Discussion of "The Unreliability of Credit-to-GDP Ratio Gaps in Real Time: Implications for Countercyclical Capital Buffers"

Simon van Norden
HEC Montréal

1. Introduction

In response to calls for banking system reforms in the wake of the global financial crisis, in 2010 the Basel Committee on Banking Supervision (BCBS) published a consultative document (BCBS 2010a) and operational guidelines (BCBS 2010b) for regulators on countercyclical capital buffers. The central feature of their proposal is the use of cyclical movements in the credit/GDP ratio to trigger increases in the reserves required to be held by banks. This in turn reflects the claim by Drehmann et al. (2010) that measured cycles in this ratio provide a useful, if imperfect, leading indicator of banking crises.

Edge and Meisenzahl (this issue) is therefore a timely analysis of a new and important change in international banking regulation, and although they consider only U.S. data, they have produced a carefully crafted empirical study. I have no substantive issues regarding the size of the revisions that they calculate, the role they find that data revision plays, or their selection of detrending methods. Their counterfactual estimates of the implied reductions in bank lending for a given change in capital requirements also seem plausible and transparent to me. I'll therefore concentrate my remarks on what I think their analysis implies for the BCBS countercyclical capital buffer framework: nothing (or almost nothing). To understand why, we should consider the analytical framework that they use as well as how the case that they examine compares to the operational guidelines in BCBS (2010b).

2. Framework

2.1 Measuring Outcomes and False Alarms

To understand the risks that credit ratio gaps will cause costly false alarms, Edge and Meisenzahl apply the methodology of Orphanides and van Norden (2002), which measures the degree to which a policymaker's measures of gaps are revised ex post. Their logic for doing so is as follows:[1]

(i) BCBS (2010a) associates excessive credit growth with increased systemic risk.
(ii) BCBS (2010a) argues that excessive credit growth can be well captured by deviations from trend.
(iii) The best measures of deviations from trend are ex post.

However, the "trends" defined in point (ii) are not ones best measured ex post, as claimed in point (iii). Rather, BCBS (2010b) and Drehmann et al. (2010) simply argue that "ex ante" or one-sided deviations from trend can be useful leading indicators of systemwide risk. (More precisely, BCBS 2010b, p. 3, states that "buffer decisions should be guided by the objectives to be achieved by the buffer, namely to protect the banking system against potential future losses when *excess credit growth is associated with an increase in system-wide risk* [italics added]." It further states [p. 3] that "the credit/GDP guide is a useful common reference point in taking buffer decisions. . . . [g]iven the guide's close links to the objectives of the buffer and its demonstrated usefulness in many jurisdictions as an indicator of the build up of system-wide risk in a financial system in the past.") They show no interest in ex post measures of "credit gaps," preferring to use episodes of banking crises or serious distress identified in earlier research by a variety of authors as their target.

Furthermore, the analysis presented in Drehmann et al. (2010) seems to show that, on average, ex ante credit-gap estimates usefully identify periods of increased systemic financial risk. Edge and Meisenzahl do not dispute this claim, not even in the specific case

[1]See Edge and Meisenzahl (this issue), second-to-last paragraph in their Introduction.

of the United States. There is no evidence that ex post measures of the credit/GDP gap are better predictors of systemwide risk than the one-sided measures. For these reasons, I do not understand why Edge and Meisenzahl feel that *the revision of the gap* is a good indicator of false alarms. At the very least, they could compare its results to the data on false alarms provided by BCBS (2010b) and Drehmann et al. (2010). Drehmann et al. (2010, esp. pp. 15–18) have an extensive discussion of the degree of type I (missed crises) and type II (false alarms) errors across various potential indicators and look to minimize the degree of type II error subject to a maximum limit on type I errors. They also argue that the costs of type I errors are likely to be higher than those of type II, and that many type II errors are classed as such only because the signal for the crisis was received too early!

The extent to which Edge and Meisenzahl improve on the analysis of one-sided gaps is slight; they examine the role of data revisions for one country and conclude that it has no significant effect. At the margin, I think this supports rather than undermines the use of the proposed Basel framework.

3. Operational Guidelines

While the credit ratio may help to predict periods of systemic financial risk, there is no question that it is an imperfect predictor. BCBS (2010b, p. 3) notes that the ratio serves only as a guide, and "the guide does not always work well in all jurisdictions at all times. Judgment coupled with proper communications is thus an integral part of the regime."

The BCBS further emphasizes that the threshold at which to increase the capital buffer and the amount of the increase need to be tailored to each country's particular experience.[2] However, Edge and Meisenzahl make no attempt to determine which threshold values work well; they simply assume that once the gap reaches the 90th percentile of its historical distribution, capital requirements are discretely increased by between 0.5 percent and 2 percent. We have no

[2]They go much farther than this, noting that countries may also choose to develop additional indicators.

information on whether other settings of credit-gap threshold could reduce costs (or costs relative to benefits), but such analysis would seem to be part of the judgment that is expected in the deployment of the countercyclical capital buffer.

It should also be noted that BCBS (2010b) proposes a different "default" approach to the deployment of the buffer.[3] It suggests that authorities define two trigger levels $\{L, H\}$, $L < H$. L defines the level of the gap at which the buffer begins to be deployed, while H defines the level above which the buffer is set at its maximum (2.0 or 2.5 percent).[4] In between, the amount of reserves to be kept as a buffer should increase linearly with the gap. This means, for example, that during the 2001 and 2003 episodes studied by Edge and Meisenzahl, the increase in reserves is likely to be smaller than it would be during more severe episodes such as the 2006–08 period. If true, this implies that their calculations overstate the degree to which bank lending would be constrained during the two periods they examine.[5]

To understand the likely costs of the capital buffers proposal, I found Edge and Meisenzahl's citation of Hanson, Kashyap, and Stein (2011) to be particularly interesting. The latter argue that during most periods of systemic financial strain, markets are willing to fund only the most heavily capitalized banks, thereby forcing increases in bank capital requirements that may exceed the maximum size of the proposed countercyclical capital buffers.[6] This suggests that if (i) the countercyclical buffer framework generates more correct deployments than false alarms, and (ii) its operation avoids even

[3] See BCBS (2010b, pp. 13–14).

[4] The BCBS suggests using $L = 2\%$ and $H = 10\%$ as a benchmark.

[5] Using the BCBS's benchmark values of $\{L, H\}$ would also imply that the buffer would have begun deployment several years before the 2001 recession and remained in place until the most recent crisis. This would probably further attenuate the degree to which bank lending would have been constrained during the downturn.

[6] In an earlier draft of their paper, Edge and Meisenzahl showed that capital ratios among the largest banks increased by roughly 400 basis points in the wake of the 1991 credit crunch and by roughly another 300 basis points in the wake of the 2008 crisis. The BCBS (2010b) framework does not contemplate changes beyond 250 basis points.

larger credit contractions caused by market crises, then the benefits of its deployment could outweigh its costs.[7]

References

Basel Committee on Banking Supervision. 2010a. *Consultative Document: Countercyclical Capital Buffer Proposal.* Basel, Switzerland: Bank for International Settlements.

———. 2010b. *Guidance for National Authorities Operating the Countercyclical Capital Buffer.* Basel, Switzerland: Bank for International Settlements.

Drehmann, M., C. Borio, L. Gambacorta, G. Jiménez, and C. Trucharte. 2010. "Countercyclical Capital Buffers: Exploring Options." BIS Working Paper No. 317.

Hanson, S., A. Kashyap, and J. Stein. 2011. "A Macroprudential Approach to Financial Regulation." *Journal of Economic Perspectives* 25 (1): 3–28.

Orphanides, A., and S. van Norden. 2002. "The Unreliability of Output-Gap Estimates in Real Time." *Review of Economics and Statistics* 84 (4): 569–83.

[7] An important question in this regard is whether the relaxation of the buffers in the midst of a crisis would lead to any new lending if the market requires capital ratios in excess of the regulatory requirements.