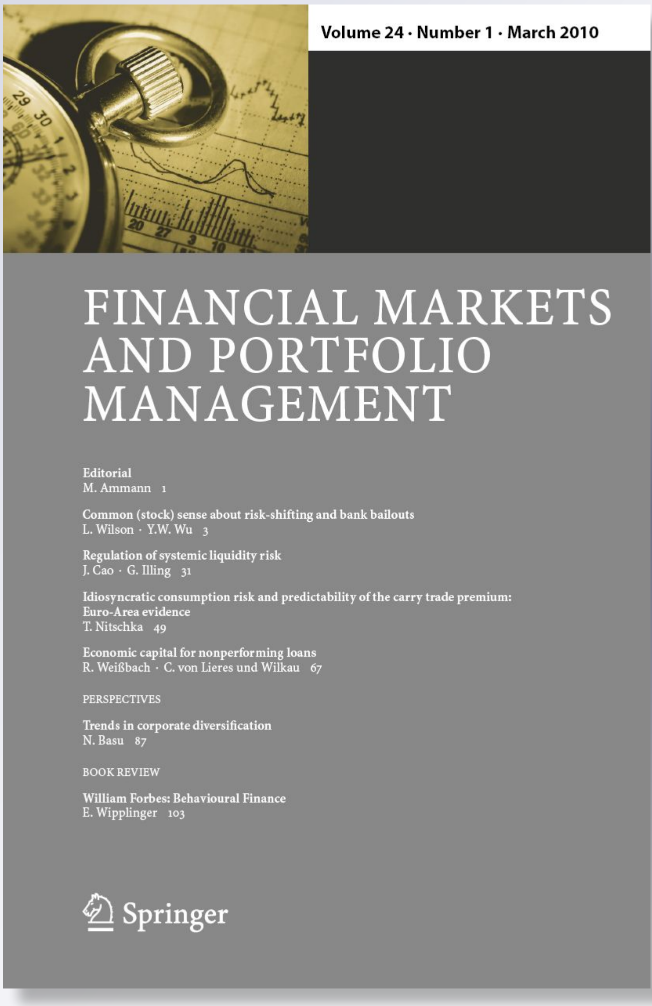
Financial frictions and real implications of macroprudential policies

Alexis Derviz

Financial Markets and Portfolio Management

ISSN 1555-4961
Volume 26
Number 3

Financ Mark Portf Manag (2012)
26:333-368
DOI 10.1007/s11408-012-0189-y



Springer

Financ Mark Portf Manag (2012) 26:333–368
DOI 10.1007/s11408-012-0189-y

Financial frictions and real implications of macroprudential policies

Alexis Derviz

Published online: 25 July 2012

Abstract I model a number of imperfections in financial intermediation that have implications for real economic activity in a production economy with technological risk. Partially opaque firms are financed by both debt and insider equity. Banks have market power over borrowers. There can be a prior bias in public beliefs about aggregate productivity (business sentiment). I investigate the dependence of equilibrium on the biased business sentiment and a prudential policy instrument (a convex dependence of bank capital requirements on the quantity of uncollateralized credit). Loss given default can be reduced by both a monetary restriction and a macroprudential restriction. Real implications of both are very similar in the aggregate, but macroprudential policies are more advantageous for bank earnings. On the other hand, the policies considered here are unable to reduce the number of defaulting firms (default frequency). Economic activity is highly sensitive to "leaning against the wind" actions on both fronts, so that using a macroprudential instrument to intervene against an asset price bubble has tangible welfare costs comparable to those of a monetary restriction. The costs can be offset by fine tuning capital charges as a function of corporate governance on the borrower side (specifically, by discouraging limited liability of borrowing firm managers).

Keywords Debt · Equity · Bank · Default · Macroprudential policy

JEL Classification E44 · G20 · G01 · E58 · E23 · D82

A. Derviz (✉)
Monetary and Statistics Department, Czech National Bank, Na Prikope 28,
115 03 Praha 1, Czech Republic
e-mail: alexis.derviz@cnb.cz; aderviz@utia.cas.cz

A. Derviz
Institute of Information Theory and Automation, Pod vodarenskou vezi 4,
182 08 Praha 8, Czech Republic

1 Introduction

Designing instruments capable of combating financial instability is high on policy-maker priority lists. The obvious reason for this is the recent global financial crisis, which provided a painful example of how an originally purely financial problem can evolve into a severe worldwide recession. For macro- and financial economists, one challenge of designing such instruments is convincingly explaining the real implications of financial frictions. Such friction would be irrelevant in a "Modigliani–Miller" world of competitive and efficient financial intermediation, a world often used as a convenient shortcut in past generations of macro models. Until recently, the bulk of macroeconomic literature posited orderly financial market operation, paying limited attention to improperly functioning financial intermediaries as a source of financial shock.

Our theoretical knowledge of financial frictions is still not fully compatible with the conventional macroeconomic modeling paradigm. The theory of financial intermediation describes many phenomena (such as informational asymmetry, agency, imperfect competition, institutional design, etc.) that impair efficient pricing of household and corporate liabilities, thus suggesting probable channels of financial friction spillover into the real economy. However, there is a certain deficit of theoretical contributions that compare the relative economic importance of various financial frictions in a setting relevant to macro theorists.

Currently, there is no general equilibrium toolkit for deriving the quantitative impact of (macro-)prudential policies intended to prevent financial frictions from reaching the threshold of financial instability. Specifically, when (some) economic agents are leveraged and credit is risky, do the policies intended to encourage prudent behavior affect the root causes or just the superficial symptoms of a nascent bubble in asset markets?

This paper begins to answer this question by proposing a model of imperfect financial intermediaries in a production economy. I consider a joint debt–equity market equilibrium with (partially insider) equity and debt financing of physical capital. For this economy, I investigate the role played by lender–borrower informational disparity, economic sentiment, and macroprudential intervention in the probability of default (PD), losses given default (LGD), and aggregate economic activity, that is, investment and output. I discuss a two-period setup, mainly for reasons of space economy; a multi-period generalization would not present a conceptual problem.

1.1 Method and main findings

For a macro model of financial frictions having the goal of addressing the matter of financial (in)stability, it is desirable to include those price variables for the major asset markets that are instrumental in channeling savings into investment, and a mechanism for their discovery. This is a natural way to adequately capture financial shock absorption by the real economy and prepare the foundation for multi-period collateral cycle modeling. It is particularly important to be able to model pricing of physical capital in its role as the commodity underlying liabilities, such as equity, debt, and collateral, created to finance production. Macro models without this capability, which includes the majority of dynamic stochastic general equilibrium with financial sector (hereinafter DSGEwFS) models designed to replicate the observed influence of financial frictions on output, have to rely on the formal workings of slack versus binding finance constraints in individual consumption and investment optimization problems. However, such a framework falls short of describing certain vital ways that financial stability contributes to economic performance. This is because autonomous price adjustment is a powerful—and nearly indispensable—way of accounting for the so-called cross-sectional side (Borio 2003) of financial turmoil. Real option theory shows some potential in this regard, but any real progress will require overcoming the currently stifling limitations of the complete market paradigm. My model departs from the DSGEwFS, reduced-form financial intermediation, and the real option literature in that it places incomplete markets with producer liabilities directly in the center of analysis. (Here, I concentrate on two: firm debt and equity. In the chosen two-period setup, markets for collateral are necessarily rudimentary. More sophisticated modeling of these markets capable of explicitly representing the collateral cycle phenomenon would be important in a dynamic extension of the model.)

Despite the reduced time dimension and the several methodological innovations in modeling credit supply and demand that I introduce in comparison to the existing DSGEwFS models, this paper has key conceptual links with the DSGE literature which influences my choice of analytical tools. DSGE modelers long ago abandoned the search for closed-form solutions in light of the fact that pursuing them would trivialize their constructions to the point of near meaninglessness: joint optimization of investment, production, and consumption simply cannot be adequately represented within the limits of high-school algebra.

Today, the best-known DSGE research adopts from its deterministic prototypes (neoclassical and neo-Keynesian growth models in particular) the convention of explaining its results by means of impulse–response exercises. This widely accepted device is applied after the model equations have been log-linearized around a "nonstochastic steady state." However, this technical trick bears the risk of misleading interpretations (a well-defined stochastic dynamic generically converges to a limit random variable, not a point, and the nonstochastic steady state is rarely a sufficient statistic even for the mean of the limit distribution, let alone higher moments). It also obfuscates the central role of higher moments of exogenous uncertainty distributions for the properties of model dynamics.

As a consequence, complete transition probabilities and limit distributions, instead of artificial nonstochastic steady states, are now popular elements of DSGE models. In the DSGEwFS field most closely related to the present paper, recent examples of this trend include He and Krishnamurthy (2010) and Mendoza (2010).

This paper takes a shortcut, compared to the "full distribution method," in that I abstain from discussing convergence dynamics, instead working with a two-period model that could be, at the expense of substantially more complex calculations, extended to an infinite-period one. I take this approach in the firm belief that (1) it is a useful first step in studying the complicated dynamic phenomena that can arise in a full-fledged DSGE environment, (2) major effects of interest can be detected in this reduced set-up, and (3) any qualitative features revealed by this investigation will more or less carry over to a meaningful multi-period generalization. In addition, by ignoring, for the moment, issues relevant for the transitional dynamics, I can concentrate on a few selected features of the limit distribution of a small-dimensional state process. Specifically, in addition to credit volume, debt, and equity prices, I look at mean output as well as two key higher-moment loan portfolio statistics of interest for financial regulation: LGD and probability of default. Although the model's reduction of the time dimension from infinity to two does not eliminate the need for numerical solutions, it does allow considerable economy in terms of both the number of variables and notation (to say nothing of computational requirements). So, when I present my findings later in the paper, I ask that the reader keep in mind conventions familiar from financial macroeconomics (i.e., mostly based on numerics) rather than expecting bona fide financial intermediation models (which are traditionally basic enough to allow for closed-form solutions).

Among the advantages of the defined environment is the possibility of linking public sentiment, so often invoked when asset price misalignments are discussed, with real investment decisions and output. Specifically, I analyze the role of incorrect sentiment about the distribution of total factor productivity across borrowers. Sentiment is identified with biased prior public beliefs. Firms can be maximally truthful in that they send unbiased signals to the public about their average productivity level. However, each firm's public signal is noisy because the firm-specific productivity type and the aggregate disturbance cannot be reported separately. Therefore, updating a prior belief may reduce the bias somewhat, but will never completely eliminate it. In such a situation, I say that there is a (prejudiced) public sentiment. Not surprisingly, prior prejudice impacts equilibrium equity prices, lending rates, bank credit, investment volume, and output (hereafter referred to collectively as "economic activity"). In my two-period setup, this belief-share demand connection can be considered as the nucleus of an endogenous equity bubble. Even in two periods, my model includes basic elements of the credit cycle effect as optimistic beliefs support inflated equity prices that, in the role of collateral, inflate credit demand, and so forth. A proper multi-period extension of this construction would have the familiar attributes of self-fulfilling price expectations and a sudden correction the moment improved information becomes available. Note that I use the term "bubble" in the present two-period model even though "sentiment-driven risk mispricing" would be more accurate. This should be acceptable shorthand as long as it is recognized that, regardless of the number of periods in the model, the term bubble is generally understood as a self-validating asset price distortion due to nonfundamental factors.

The main application considered in the paper is the introduction of a policy tool designed to counteract systemic credit risk. Specifically, I look at the impact of additional (and convexly growing) regulatory capital charges on banks that lend to firms with low relative equity. Although the true advantages and disadvantages of such policy instruments are fully revealed only in a dynamic model, I am nevertheless able to gauge the basic qualitative consequences of said policy for economic fundamentals within each period. This is made possible by uniting the features of a usual model of production with financial friction effects.

We will deal with three types of friction. The first type involves imperfectly substitutable equity and debt markets for risky and opaque producers. Such markets become a source of nondiversifiable systemic risk. The second friction arises from banks that have a nonzero market power over borrowers. The possibility of biased public beliefs/prejudiced sentiment as a source of asset bubbles constitutes the third type of friction. To the best of my knowledge, a synthetic analysis of all three phenomena in a common model is novel to the financial macroeconomics literature. The model allows me, among other things, to investigate real economic implications of macroprudential policies motivated by financial stability considerations.

To briefly preview my results, I find that

- macroprudential capital surcharges on banks are successful in reducing LGD, and marginally alleviate the real economy implications of correcting biased public sentiment.
- however, macroprudential tightening fails to reduce PD and leads to a contraction of economic activity of the same order as the LGD reduction.
- macroprudential tightening results in a downward pressure on equity (physical capital) prices and a modest reduction of leverage; this result might explain why policymakers find macroprudential restrictions such an attractive strategy in dealing with credit bubbles; however, output losses are a warning that the costs may exceed benefit. In short, macroprudential policy is more successful at alleviating the symptoms of a bubble than at curing the underlying disease.
- The aggregate implications of monetary and macroprudential policies aimed at bursting the same asset price bubble are strongly aligned and may be difficult to separate empirically; however, *ceteris paribus*, macroprudential tightening is more advantageous for banks.

1.2 Related literature

In the past, investigations into the interplay between financial and real shocks at the macro level were fairly rare. The concept of costly state verification (CSV) in contract theory (Townsend 1979) has gradually found its way into real business cycle models (initially by suggesting an appropriate way of modeling default on debt contracts). Inspired by CSV models, the financial accelerator construction of Bernanke et al. (1999) became an influential example of feeding a financial sector factor into quantitative macroeconomic theory. However, the stated and met objective of Bernanke et al. (1999) was to *codify*, not necessarily *explain*, the main implications of financial sector presence in the economy, as the authors strived to reflect empirically important business cycle phenomena related to financial frictions. In essence, Bernanke et al. (1999) and the succeeding DSGEwFS models (e.g., Christiano et al. 2008) obtain financial frictions by including debt-to-net worth, liquidity, and other similar constraints in optimization problems of agents. Thus, although DSGEwFS models accommodate financial intermediaries, they usually assume proper functioning of such intermediaries. As the very term “financial accelerator” suggests, the financial sector shapes the real shock propagation mechanism in the economy, but does not itself originate the shock. This is because capital suppliers do not possess sufficient prerequisites with regard to either standing in the market or informational endowments.[1] Therefore, they are unable to "misbehave" in a natural way (e.g., in terms of adverse selection, reputation, incomplete contracts, herding behavior, etc.) along the lines drawn by financial intermediation theory. Financial intermediation theory, however, relies on toy models that usually provide very indirect empirical guidance. Another insufficiently developed link in the current state of the financial accelerator literature is the one it has with the asset pricing theory. Naturally, the latter, to the degree it is trapped in the efficient market paradigm, does not make synergies any easier.

Interest in macro theory with financial frictions soared with the outbreak of the global crisis in 2007, an interest later extended to models of regulatory policies intended to respond to a financial shock. All this attention spilled over to rekindle interest in formal modeling of macroprudential policy tools that augment standard Taylor rule-based interest rate policies. Quantitative assessments based on tentative synthetic techniques were conducted in attempts to understand the financial crisis and resulting global recession (see, e.g., Chapter III of the IMF October 2009 World Economic Outlook; the exercise undertaken there draws on the approach of Aoki et al. 2004; Iacoviello 2005; Monacelli 2009). Naturally, a proper quantitative analysis of the workings of those additional instruments requires a more explicit role by financial intermediation than was usual in earlier macro models. My paper constitutes a step in this direction, as I propose a fairly general way of introducing macroprudential instruments in a production economy with a financial sector. Unlike some other recent contributions that, although taking both corporate and bank default into consideration, omit systemic driving factors of default from the model (de Walque et al. 2010), I preserve the main features of the risky lending paradigm of the financial intermediation literature (see, e.g., Stiglitz and Weiss 1981), but connect this paradigm to those of neoclassic and neo-Keynesian production economies.

Several contributions to the literature note the impact that the marginal rate of prudential capital charges can have on the real economy. Blum and Hellwig (1995) analyze the macro consequences of capital requirements in the language of undergrad textbook macroeconomics (a traditional reduced-form Keynesian IS–LM macro model) without explicit markets for either physical capital or bank credit. Bank capital requirements, essentially, are just an ad hoc friction fed into investment demand; there is neither income uncertainty nor default risk. Compared to my model, the utilized IS–LM environment includes a richer set of variables and produces predictions for a wider range of macro phenomena related to capital requirement adjustments, for example, consumer price and wage effects, in addition to investment and output. At the time of its publication, Blum and Hellwig (1995) provided an elementary formal underpinning of the now well-accepted thesis that capital requirements are likely to serve as real shock amplifier. In a state-of-the-art modeling environment of DSGEwFS, Covas and Fujita (2010) generate a formally different picture of

[1] One example is the full competitiveness assumption, which imposes the zero-profit constraint on lenders. What may be a gain in analytic convenience (a reduction of the number of free parameters) is also a loss in flexibility, since the lender's market power is a feature one would really want to model. Moreover, it is often overlooked that zero profit is a two-way "egalitarian" constraint: not only is economic profit prohibited, but so are losses. But to model a bank without downside risk is nearly meaningless, as the recent financial crises painfully illustrated.

real business-cycle effects of bank capital restrictions. That paper embeds the intra-period agency problem between households and entrepreneurs in an otherwise standard DSGE setting. Accordingly, Covas and Fujita (2010) work with only a partial (in the trio household–bank–entrepreneur) equilibrium in the financial intermediation part of the model within each period, whereas general equilibrium mechanisms are reserved for the dynamic macroeconomic part. Thus, their model does not track the financial-real friction spillover in specific markets any more than did Blum and Hellwig 15 years earlier. Moreover, both Blum and Hellwig (1995) and Covas and Fujita (2010) posit a simple linear form of bank equity dependence on credit (Cooke ratio) that is a far cry from current macroprudential regulation. In my model, penalties for lending to overleveraged producers (who, although individually rational, are unable to endogenize aggregate excessive leverage) are introduced to fight excess defaults, which is in accord with the current consensus in the literature concerning the targets of macroprudential controls (cf. Angelini et al. 2010). Although my results concerning real consequences of macroprudential regulation can be seen as broadly consistent with the findings of Blum and Hellwig (1995), Covas and Fujita (2010), and others (e.g., Repullo and Suarez 2010), methodologically my analysis adds the important aspect of simultaneous clearing of all markets relevant for production financing.

In addition, my model predicts a high sensitivity of economic activity to the manager incentives in borrowing firms. In this regard, my results are in alignment with the literature that discusses the far-reaching implications of choice of manager incentive scheme under separation of ownership and control in DSGE models (Donaldson et al. 2009).

The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 defines a parametric version of the model, with which I conduct numerical experiments; this section illustrates the role played by biased economic sentiment in asset prices and economic activity. In Sect. 4, I introduce a macroprudential policy instrument and, in parallel, experiment with modifying borrower downside risk at default. Section 5 concludes.

2 Model

2.1 Environment

There is a set $\mathbf{L}$ of productive industries in which agents can invest. Each industry is operated by a representative firm differing from others by the value L of its idiosyncratic total factor productivity component. Thus, elements of $\mathbf{L}$ can be identified with distinct L values standing for firm types. Each production capacity has a constant-returns-to-scale production function (to be described later), which transforms inputs provided in period 1 into stochastic revenue in period 2. All inputs, investment, and output are expressed in terms of a single unit of account. There are two periods and three groups of agents: retail investors, firm managers (or, more simply, firms), and banks.

The return on each firm's equity depends on inputs in the production function. Of those, one is physical capital and the other is managerial human capital, which, as compared to physical capital, is firm specific and, consequently, individuals cannot distribute their human capital between firms. There is one representative retail investor who is specialized in a particular industry, that is, the representative firm. For this firm, the investor is the exclusive supplier of human capital and also knows details of the firm's inputs into the production function. As a result, this "insider" investor is sufficiently "in the know" to make well-informed choices of what stock to own within the industry, but not outside it.[2] We posit an implicit mechanism akin to Townsend (1979) CSV one, but tied to the knowledge of the firm's input–output structure, that only allows the insider retail investor to discriminate sufficiently between earnings under different states of nature. Due to the same mechanism, banks and all other ("outside") retail investors who are less well-informed do not hold equity in that firm (consequently, banks do not hold any equity at all). Thus, share demand is manifested only by a subset of knowledgeable investors. Limited participation precludes full diversification of equity holdings by retail investors. Due to the limited participation constraint, nondiversifiable systemic risk can emerge, a situation that the regulatory authority may want to counteract (cf. Sect. 4).

With too few equity holders for each firm, producers also seek debt financing, whereas retail investors demand risk-free deposits (in our model, deposit-collecting institutions do not default) as a complement to risky equity. Both demands are met by banks who accept deposits with the declared objective of optimally investing them in the whole spectrum of available firms, that is, the banks act to diversify retail investor funds. As opposed to retail investors, the representative bank (here, *wholesale bank*) can lend to any firm, which is why a wholesale bank can act as a diversifying intermediary, while at the same time extracting rents as the exclusive operator of necessary financial technology and enjoying market power. Accordingly, in contrast to long-standing tradition in DSGEwFS modeling, we do not have to impose the competitive zero-profit constraint on banks (cf. Sect. 1.2, note 1).

We shall consider two types of banks: wholesale banks and relationship banks. Relationship bankers are loan officers who have better information about borrowers than either the public or the wholesale bank. More precisely, they are able to find out the borrowing firm's type, that is, the firm's idiosyncratic productivity component. This ability defines their nontransferrable human capital. There are two basic variants of the model: in one, relationship banks are active (RB economy); in the other, they are not (arm's-length lending, or AL economy). In an RB economy, the wholesale bank delegates lending to better-informed loan officers; in an AL economy, it does not. I introduce this distinction so as to be able to confront my model with some established propositions of financial intermediation theory. In the latter, the concept of banks as delegated loan monitors (Diamond 1984; Diamond and Rajan 2000) hinges on their superior ability to collect information on borrowers. This is roughly the same ability as the one we give lenders in the relationship banking regime, which, conceptually, should be considered the baseline bank operation mode. However, recent inquiry into the securitization of bank assets has revealed a substantial number of cases in which the

[2] One can imagine a household of two, with one member supplying equity financing and the other member supplying human capital input to the same family industry.

said informational advantage is voluntarily given up by loan originators. This is why I conduct the same formal exercise under the alternative of no superior lender information (our arm's-length regime). For instance, distinguishing between lending regimes is a useful way of assessing the effect of both asset bubbles and macroprudential policies under varying price discrimination in the credit market. A very crude sensitivity test can be conducted by pitting one-sided against two-sided observability of differences in productivity type. However, lending technology choice is closely related to the issue of bargaining between the wholesale and relationship bankers, an aspect I do not model explicitly. Therefore, I conduct only a comparative statics analysis of that issue here.

The financial sector defined in the model shall reflect some prominent features of the recent crisis. To this end, I assume that the common cause behind very diverse manifestations of the global financial crisis can basically be seen as the well-known agency problem of fund diversion. A financial institution sells its services to the public (e.g., collects deposits) by declaring one investment strategy, but the actual investment of the borrowed funds is different, and also such that cannot be easily observed or defined. In particular, returns on lending to an individual firm will be higher, the greater the information known about that firm, thus making it attractive for the wholesale bank to delegate investment decisions to the better-informed relationship banker. However, the relationship banker, being in possession of exclusive firm-specific knowledge, has considerable bargaining power vis-à-vis the wholesale bank. Therefore, the relationship banker can demand compensation for services up to an amount just below the point where the wholesale bank becomes indifferent between employing the relationship banker's services and managing the loan itself based only on public information. Past this point, delegation will not occur, and the loan is managed by the wholesale bank at arm's length (subsequently, those loans can be packaged, tranched, and sold to other banks in CDO form). It is also possible that, under a particular sentiment and other exogenous parameter values, there is an equilibrium with delegated loan management but no equilibrium with arm's-length management. In such a case, if wholesale–retail banker bargaining over fees is unsuccessful, a large number of firms will cease to operate because of their inability to finance production with either equity or debt, and output will be considerably reduced.[3]

Under other circumstances, an AL lending regime can be preferable both subjectively (the wholesale bank earns more on its loan portfolio) and socially (lower interest rates on average, more investment of debt-financed capital, and, hence, higher output). The problem is that the outcome is sensitive to, among other things, the quality of public information (cf. Sect. 3).

Uncertainty and information: the total factor productivity of the firms, denoted by A, is the only source of risk in the model. We assume that productivity is a product of two components: $A = LS$, where S is the aggregate uncertainty, whose distribution is

[3] This seems a plausible reflection of at least some instances of the transition from "purely financial" revision of beliefs, and the corresponding turbulence in asset markets, to the abrupt adverse impact on investment and GDP as observed during the latest crisis.

known to everyone (for simplicity, let it be the same distribution for all firms).[4] L is the firm-specific component, about which the "public," comprised of retail investors and the wholesale bank, only knows the distribution across the set $\mathbf{L}$ of industries. There are three groups of agents that have information superior to known by the public: relationship bankers, firm managers, and retail investors within their own industry. A relationship banker knows the type of each firm that applies for a loan. The manager of that firm also knows its type. Managers also know inputs in the production function (seeing as it is they who decide on them), but we assume that they are contractually forbidden to trade in their firm's stock. Accordingly, note that retail investors do not know the productivity type of other firms, not even those in the same industry. This type of knowledge is different from their insider knowledge of technology and is known only to managers and relationship bankers, thus justifying separate positions of these two agent types in the model. So, the information known to the inside investor is inferior to that known by the manager, qualifying the latter to negotiate a loan with the bank.

Finally, the bank approached by a potential borrower knows the functional form of its loan demand depending on the lending rate. Under the AL lending regime, this demand function contains a distributed parameter (type L; for more detail, see Sect. 2.4).

The above implies that all agents in the model suffer from certain knowledge limitations. Retail investors buying shares can observe the lending rate, but are unable to infer the productivity type. Similarly, a bank setting an interest rate knows the equity price of the borrower and anticipates the functional form of its loan demand, but is not sophisticated enough to infer productivity type from the latter. Accordingly, the possibility of a separating equilibrium is not an issue even in the special case of our simple model with only two productivity types, with which we deal in simulations in Sects. 3 and 4. This "black-box" nature of perception can be justified by the potential presence of unobservable parameters that prevent full identification. (For example, models with continuous support for random factor distributions typically do not allow for such identification.) This is a common, if unspoken, convention of rational expectations macro models.

The objectives and choices of the defined agents are described next.

2.2 Retail investors

Each retail investor has a stock of initial wealth w_0 and a stock m_l of nontransferable human capital in exactly one industry $l \in \mathbf{L}$. This human capital is sold to firm l in period 1 at price z^l. For simplicity, we assume that the supply of human capital is inelastic, that is, the whole stock m_l is sold regardless of the value of z^l. This same investor, or the second member of the same household, can use cash $w_0 + z^l m_l$ available in period 1 to either buy shares in firms of the same industry l or put the money

[4] One can compare this feature with Bernanke et al. (1999) and successor models: these, too, contain both aggregate and firm-specific uncertainty, but the role of the former is played down, with substantial impact on interpretation of results. Indeed, when systemic uncertainty is present, Bernanke et al. (1999) do not even have a proper debt contract in the model, and the state-contingent hybrid they use instead is difficult to rationalize. On the contrary, my model includes systemic uncertainty as a key fundamental factor and lets it play a role in both equity and debt pricing.

in a bank account offering a fixed interest rate i. One share earns $y^l(A^l)$, where A^l is the total factor productivity parameter. Exact expressions are set out in Sect. 2.3. The important point is that since another member of the same household supplies firm-specific human capital to l, the retail investor household knows the exact levels of inputs in the production function. Therefore, even though productivity realization in period 2 is uncertain, the degree of uncertainty is much lower than it would be if the investor decided to buy shares in another industry $n \in \mathbf{L}$. For an outsider, only return y^n without a breakdown into factor inputs and productivity would be known, which would combine the uncertainties over A^n, physical capital k_n (see Sect. 2.3), and m_n. Without going into technical detail, we assume that the resulting uncertainty is so high that it is never optimal for any retail investor to reduce share holdings in the inside industry and buy shares in outside ones.

If the investor buys x_l shares in industry l at price p^l (which he takes as given), his wealth in period 2 is equal to

$$w = x_l y^l + (1+i)(w_0 + z^l m_l - p^l x_l).$$

This final wealth, which is uncertain due to the uncertainty in A^l, enters the investor's utility function, whose conditional expectation in period 1 is maximized with respect to permissible choices of x_l. The interval of permissible choices is [0, 1]. This means that the number of shares in each industry is normalized to unity and short selling is not allowed.

Denote the investor utility by U and his subjective beliefs about the distribution of A^l values by ϕ. We consider only continuous nonatomic distributions so that ϕ is a well-defined density. Then the investor solves the problem

$$\sup_{0 \le x_l \le 1} \int U(x_l y^l(A) + (1+i)(w_0 + z^l m_l - p^l x_l))\varphi(A)\,\mathrm{d}A. \tag{1}$$

The outcome can be either an internal solution characterized by the first-order condition

$$\int U'(x_l y^l(A) + (1+i)(w_0 + z^l m_l - p^l x_l)) \left[y^l(A) - (1+i)p^l \right] \varphi(A)\,\mathrm{d}A = 0 \tag{2}$$

or a corner solution in situations where the left-hand side of Eq. (2) does not change sign for $x_l \in (0, 1)$. We exclude from consideration the trivial corner solution $x_l = 0$ (firms without any outside equity capital) and consider the remaining cases.

The internal solution can be characterized by a conventional finance theory formula. Under market clearing (the representative investor holds $x_l = 1$), it can be restated as

$$p^l = \frac{1}{1+i} \int \frac{U'(y^l(A) + (1+i)(w_0 + z^l m_l - p^l))}{M(z^l m_l, p^l)} y^l(A)\varphi(A)\,\mathrm{d}A, \tag{3}$$

with $M(h, p) = \int U'(y(A) + (1+i)(w_0 + h - p))\varphi(A)\,\mathrm{d}A$.

Equation (3) can be interpreted as the expected payout on firm l stock discounted by the subjective stochastic discount factor. The value of the latter under productivity realization A is equal to $1/(1+i)$ times the investor's marginal utility of wealth U' under A, normalized by expected marginal utility M.[5] But whereas standard asset pricing theories concentrate on the market pricing of risk that follows from the properties of the stochastic discount factor, we note that the right-hand side of Eq. (3) also depends on p^l and thus view Eq. (3) as an equation that determines this price implicitly.[6]

Additionally of interest is the corner solution $x_l = 1$, which obtains when the objective function [Eq. (1)] of the representative retail investor is increasing in x_l on the whole interval (0, 1). Equivalently, the left-hand side of Eq. (2) is always positive in x_l and the investor actually is able to buy all the available stock for less than the expected discounted payout:

$$p^l < \frac{1}{1+i}\int \frac{U'(y^l(A)+(1+i)(w_0+z^l m_l - p^l))}{M(z^l m_l, p^l)} y^l(A)\varphi(A)\,\mathrm{d}A. \tag{3C}$$

It is quite possible that there will be a whole continuum of prices satisfying this inequality, giving rise to multiple equilibria—an additional source of potential volatility not just in asset prices, but also in interest rates, investment levels, and output. Under some parameterizations of the model, a switch from a unique equilibrium implied by the internal price solution [Eq. (3)] to equilibrium multiplicity corresponding to a continuum of corner price solutions [Eq. (3C), C for "corner"] is possible by a mere shift of sentiment (a formal definition and an extended discussion of the latter can be found in Sect. 3).

2.3 Firms

2.3.1 Baseline: unlimited liability of producers

Firms transform physical capital k and human capital m into output by means of c.r.s. production functions. Although our main variable of interest is physical capital, the second input allows us to equip the model with both a liquidity constraint on the borrower side and a source of leverage. The latter emerges because (cf. Sect. 2.2) the sum of zm across retail investors acts both as the cash they deposited in banks (in excess of the initial wealth) and the lower bound of credit granted by banks to firms. Accordingly, the model generates endogenous leverage as a result of varying retail investor and bank balance sheet sizes. This makes it possible to investigate the effects of various policy measures on private-sector leverage.

Production functions contain uncertain total factor productivity. This uncertainty has two components. The first is a systemic risk factor whose distribution function is

[5] In model extensions containing the retail investor's consumption in period 1, M should be equal to the marginal utility of consumption in that period.

[6] Note that being an equation that generalizes the conventional asset-pricing formula, Eq. (3) introduces an equity-market-based (co-)determination mechanism for physical capital. Such a mechanism is absent from existing financial accelerator models.

known to everyone. The second component is firm specific (the firm's type), which is known to firm management but cannot be precisely and credibly communicated to either equity investors or wholesale banks. The firm can send a public signal as to its productivity level only as a whole, that is, it cannot send separate signals about each component. In this signal, systemic uncertainty contaminates the message about the idiosyncratic productivity component value. Only a loan manager with specific expertise (a retail relationship banker) has the necessary skills to learn the borrowing firm's type.[7] In RB economies, such a manager is hired by the wholesale bank and delegated the authority to set the lending rate and collect the proceeds.

The firm's internal funds are insufficient to cover production costs, so it seeks external financing in both equity and debt form. The firm is a price-taker in both these markets. Recall that equity is sold to a subset of retail investors (those who observe the human capital input in the firm), whereas debt financing is reserved to banks. Incorporating the experience of CSV modeling (Townsend 1979), we assume that even delegated loan managers of relationship banks are unable to observe the human capital input with enough precision to support a state-contingent (equity) contract. This allows us to exclude from consideration the case of banks holding equity.

In what follows, except when necessary to avoid ambiguity, we omit the industry index l when discussing a firm's actions.

Human capital input m must be paid for upfront in period 1. For simplicity, we assume that firms do not have sufficient initial cash holdings to do this. So a firm using m units of human capital needs to borrow from banks at least the amount zm. More borrowing may be needed to finance physical capital, for which the identity $k = k_0 + p + b$ holds. Here, k_0 is the initial nontraded "foundation" stock, for example, stock held by the company founders, p is the "market capitalization," that is, the value of shares sold in the equity market (recall that the number of shares is normalized to unity), and b is the physical capital financed by a bank loan.

In the second period, the firm produces $Af(k, m)$ units of output. We assume that the whole stock of physical capital is then released as a part of firm earnings so that, in total, they are equal to $Af(k, m) + k$. (Since this is a two-period model, it makes little sense to consider capital depreciation.) Recall that in period 2, the m input has already been paid for out of the bank loan.[8] So, the dividend to stockholders is equal to what remains of the output after the total debt, that is, $zm + b$, is serviced. Default occurs if output is not enough to repay the debt, in which case the bank seizes all earnings. Let the lending rate be r (taken by the firm as given; see more in the next subsection).

7 For simplicity, we consider only the case in which the loan manager knows the type precisely, that is, the loan manager has the same information as does firm management. Generalizations allowing the relationship banker to have "noisy" information, that is, information not quite as good as management's but better than that possessed by the public, are possible but do not add much to the qualitative insights of the model.

8 We thereby avoid the need to account for the consequences of possible firm default on the payment to m suppliers. In principle, we could have defined a contract with m suppliers receiving payment in period 2. Then, under default, these claimholders would have been pooled with the lending bank for the purposes of debt resolution. However, this would have meant unnecessary technical complications without much of a benefit to the main goal of the present analysis, which is to explore the real consequences of interactions between firms and banks.

Formally, shareholder dividends are

$$y(A) = \max\{Af(k,m) + k - (1+r)(zm+b), 0\}.$$

If the firm does not default, this dividend can also be written as

$$Af(q+b,m) + q - (1+r)zm - rb, \tag{4}$$

where $q = k_0 + p$ is total equity capital (traded and nontraded).[9]

Each firm is run by a risk-neutral manager. For simplicity, we assume that he acts in the best interest of the shareholders (i.e., we ignore agency effects in the shareholder–firm manager relationship). Then, the manager's objective is to maximize the expected dividend given the manager's superior knowledge of productivity, that is, his knowledge of firm type.[10]

Since there are exactly as many firms (industries) as there are productivity types, our use of the same letter to index the firm set $\mathbf{L}$ (lowercase l) and firm-specific productivity value (uppercase L, and lowercase $l$for its log) should not cause confusion.

The firm takes the offered lending rate r and the $m$price z as given. It is natural to assume that the equity price $p$and the overall level of equity capital q are also exogenous to the firm. Optimal levels of m and b are chosen with the knowledge that, in default, the dividend is zero. The critical level of systemic production uncertainty above/below which the firm survives/defaults is[11]

$$S^d = \frac{(1+r)zm + rb - q}{Lf(q+b,m)}. \tag{5}$$

Therefore, the firm's dividend expectation is calculated over realizations of S exceeding S^d. Let us denote the cumulative distribution function of S by X and the corresponding density by χ. The survival probability is then $X^+(S^d) = 1 - X(S^d)$, and we will also need the notation

$$\Psi^+(S^d) = \int_{S^s}^{+\infty} S\chi(S),\ \bar{S} = \int_0^{+\infty} S\chi(S),\ \Psi^-(S^d) = \bar{S} - \Psi^+(S^d), \quad \theta(S^d) = \frac{\Psi^+(S^d)}{X^+(S^d)}.$$

[9] Note the difference between our q variable and the net worth variable of Bernanke et al. (1999) and successors: since the latter (financial frictions) models do not have explicit equity markets, their net worth value is monolithic, whereas mine is naturally split into foundation and traded stock.

[10] Exact L knowledge by both the firm manager and the delegated loan manager (relationship banker) is a useful technical simplification, but not central to the qualitative results. What is important is that the amount of knowledge held by the firm and the relationship bank, even if different, is higher than that known by the retail investor and the wholesale bank.

[11] Although this cut-off value is formally analogous to similar parameters used by Bernanke et al. (1999), Christiano et al. (2008), and related models (the usual notation being $\bar{\omega}$), remember that my critical productivity value refers to systemic uncertainty realizations *conditional on the given firm-specific uncertainty*, whereas the mentioned papers deal with the firm-specific component.

Note that $\theta(S^d)$ is the expected systemic productivity component conditioned on survival.

Lemma 1 *Given the equity capital level q, human capital price z, and lending rate r, the optimal decisions of a firm of productivity type L on m and b are characterized by the first-order conditions*

$$\theta(S^d)Lf_m(q+b,m) = (1+r)z, \tag{6a}$$

$$\theta(S^d)Lf_k(q+b,m) = r. \tag{6b}$$

[In Eq. (6), subscripts denote partial derivatives.] The proof is straightforward given that, in the expression for expected dividends, only realizations of S that exceed S^d are integrated. As a consequence, the marginal products enter the first-order condition with the tail expectation multiplier $\Psi^+(S^d)$, whereas the remaining part of the partial derivative of the dividend expression [Eq. (5)] does so with the survival probability multiplier $X^+(S^d)$.

Since we assume a fixed supply of m, Eq. (6a) will be interpreted as a market-clearing condition for z, that is, characterization of the human capital price that equalizes the fixed supply with the demand determined by the marginal product of m. The second optimality condition, Eq. (6b), is an implicit characterization of the credit demand $b = B(r)$ as a decreasing function of the lending rate charged. This is the firm's reaction function in the game it plays with the bank (cf. the next subsection). Naturally, B also depends on q, z, and the parameters of the model, but we omit them to simplify notation.

2.3.2 Extension: producer downside risk

Since the production function is c.r.s., by combining Eqs. (5), (6), and the Euler identity, we arrive at the following condition for the survival threshold S^d:

$$S^d = \theta(S^d) - \frac{(1+r)q}{Lf(q+\hat{b}(S^d),m)}. \tag{7}$$

In Eq. (7), $\hat{b}(S^d)$ is the optimal choice of b implied by Eq. (6). The above condition is an equation for S^d whose solution depends on z, r, and q as parameters. The problem is that for typical distributions, production functions, and a subset of otherwise realistic parameter values, this equation may have either two solutions or none at all. In the latter case, equilibrium equity + debt financing of such a firm cannot exist either, regardless of the presence of other firm types in the economy. In the former case, two equilibria are possible, corresponding to high/low debt-financed levels of capital and high/low default probability in this firm type. Thus, our model is able to imitate real economic instability as a result of tiny financial shocks.

This type of instability is a direct consequence of limited liability, specifically, the fact that realizations of S implying default are not taken into account by the manager (he receives zero remuneration under any of them). If downside risk were present, there would be one less source of equilibrium multiplicity about which to care.

More precisely, assume that the firm manager maximizes the unconditional expectation of after-interest earnings (i.e., including the expectation over those Srealizations that would make net earnings negative in the absence of limited liability). Such a manager will borrow the following "unlimited liability" quantity of funds:

$$\bar{S}Lf_m(q+b,m) = (1+r)z, \tag{6aUL}$$

$$\bar{S}Lf_k(q+b,m) = r \tag{6bUL}$$

(Recall that $\bar{S}$ is the unconditional mean of systemic productivity component S.) So, although the default consequences for the lender are the same as in the limited liability case, that is, the bank seizes the output, the value of which is insufficient to repay the debt in full, the manager behaves "as if" he bore the full brunt of insolvency. To make managers behave in this fashion, one would need, for example, a compensation scheme that is a function of after-interest earnings, e.g., a fixed fee plus a percentage of actual—positive or negative—earnings. Similar "proportional liability" remuneration schemes, in a much more general setting than the present one, are considered, for instance, by Hui (2003). In general, the socially optimal manager fee is likely to be strongly dependent on the agent's informational advantage over the funds-providing principal (shareholder). Such a dependence is present even in the linear payoff case (i.e., when, instead of a concave a production function and limits to losses on the downside, as in our setting, there is a flat stochastic return per investment unit), as shown, for example, by Kraft and Korn (2008) for the continuous time portfolio optimization case. However, detailed investigation of the optimal managerial fee dependence on information asymmetry is outside the scope of the present paper.

In any event, firm choices based on Eq. (6UL) instead of on Eq. (6) lead to the following analogue of Eq. (7):

$$S^d = \bar{S} - \frac{(1+r)q}{Lf(q+b,m)}. \tag{7UL}$$

Now, the default threshold is uniquely determined by the endogenous variables b, q, r, and the parameters of the model, and the problem of equilibrium indeterminacy disappears. However, managerial compensation schemes able to induce "unlimited liability behavior" are mostly hypothetical and rarely encountered in practice. Therefore, counting on financial intermediation disruptions following from the limited liability case [Eq. (6)], Eq. (7) is an empirical necessity.

2.4 Banks

2.4.1 Lender–borrower negotiation

The lending bank's interaction with the borrower is assumed to take the form of a leader–follower game in which the bank is the leader and the firm is the follower. If a firm submits a credit application to a bank, the bank makes an interest rate take-it-or-leave-it offer and the firm decides whether to accept the offer. That is, the firm decides on an optimal reaction to every value of the proposed lending rate (reaction function) and the bank sets the lending rate based on the information it has about this reaction function.[12]

The bank's action depends on whether it is the original wholesale bank that negotiates the loan or whether negotiations are delegated to a relationship banker. In the first case, the bank has a belief distribution over the borrower's productivity value A as a whole (convolution of beliefs about S and L). In the second case, the delegated loan manager knows type L exactly (as does the firm manager) and faces aggregate uncertainty only regarding S. As a result, the wholesale banker sets a common interest rate for all borrowers, whereas delegated relationship bankers set separate rates for individual types.

The above definitions may raise a question about robustness with regard to the lender–borrower bargaining protocol, and, more generally, why we consider only the take-it-or-leave-it option for the firm. To be sure, microeconomics of financial intermediation employ models of borrowers who reveal their quality when they decide about loan volume. However, the hypothetical question of whether the AL regime (firm type unknown) would survive if the bank made a menu offer (rate–volume pair) is of minor importance to me given that in any version of the model with a continuum of types, separating equilibria would be irrelevant anyway. My preference was for a definition that would survive transplantation into a DSGE-conformed environment of large random event spaces (i.e., more than one source and continuous distributions of risk, large sets of productivity types, multiple agency problems, variability in time, etc.). A further possible specification of a simultaneous-move lender–borrower game would not admit a pure strategy equilibrium for standard concave production functions unless there were exogenous restrictions on maximum and minimum volumes and rates. If such restrictions are introduced, one ends up with a corner solution. That is, the bank–firm bargaining problem becomes moot and the credit market equilibrium is now a function of these ad hoc assumptions, for which economic rationale remains to be found, or, in other words, one is back at step one. Alternatively, an equilibrium in which a bank offers a rate–volume menu, like in Besanko and Thakor (1987), is possible in principle, but would require quite limiting assumptions as to the information precision on both sides. In an environment of the Besanko–Thakor type, with a narrow specification of exogenous uncertainties, the productivity type would need to be known, thus eliminating the AL lending possibility. On the contrary, for the environments targeted by my model (i.e., those compatible with DSGE extensions; cf. the discussion in Sect. 1.1), type identification is irrelevant, making the AL regime a legitimate outcome.

A Nash bargaining solution to the bank–borrower game is also possible, as found in an analogous situation by, for example, Diamond and Rajan (2000). Also under that

[12] This setup endows the bank with market power (cf. Sect. 1.2, note 1). The fact that, generically, a bank–client relationship is not fully competitive on either side is long recognized in the literature (cf. Santomero 1984). Moreover, as described by Saunders and Walter (2012), recent technological advances, geographical expansion of banking business opportunities, and the financial risk and policy response globalization during the latest crisis have allowed many banks with already considerable precrisis market shares to garner even more market power in its wake. Popular examples of imperfect competition modeling use the concept of client "catch-up" in a specific bank (see, e.g., Bonaccorsi di Patti and Dell'Ariccia 2004).

option, the determinants of equilibrium in the credit market would not be pinned down but instead hidden in poorly justifiable assumptions, in this case in the selected parameterization of bargaining power of the parties (the outcome would be very sensitive to the choice of such a parameterization). In my investigation of these and other alternatives for characterizing the lender–borrower interaction, I was looking, with an eye toward future embedding in a standard stochastic macroeconomic model, for the least artificial ones with regard to state space size, technologies, and information endowments. Under these criteria, the bank-moves-first option was preferred both for its adequate representation of bank market power over established borrowers (cf. Sharpe 1990; Diamond and Rajan 2000; Dell'Ariccia and Marquez 2004) and its empirical relevance (cf. the high share of credit lines among modern corporate loan contracts).

2.4.2 Lending rate determination

Banks are assumed to be risk neutral. The bank incurs a cost of funds, which, for simplicity, we denote by i (the same as the deposit rate for retail investors) and assume a linear funding price regardless of volume. In Sect. 4, we look at the consequences of relaxing the last of these three assumptions. Deviations from either of the first two assumptions can be easily accommodated in the model as well, but are less important to the context of this paper.

We formulate the rate-setting problem of the delegated loan manager first. Using the notation of the previous subsection, a firm of type L borrows $B = zm + b$, where the optimal quantities of both components are determined by the optimality conditions [Eq. (6)]. Thus, from Eq. (6a), with $\hat{m}$ and $\hat{k} = q + \hat{b}$ being the optimal levels of, respectively, human and physical capital,

$$B = \frac{\theta(S^d)Lf_m(q + \hat{b}, \hat{m})\hat{m}}{1 + r} + \hat{b}. \tag{8}$$

Since we assume that m is in fixed supply for each firm (price z equalizes this supply with optimal demand), one can drop the hat in the notation: $\hat{m} = m$. Further, $\hat{b}$ can be expressed through L, m, q, r, and $\theta = \theta(S^d)$ using Eq. (6b). Often, the expression can be made explicit. For instance, for the Cobb–Douglas production function $f(k, m) = k^\alpha m^{1-\alpha}$, the named first-order conditions imply that for optimally chosen physical and human capital, $\hat{k}$ and $\hat{m}$, $f(\hat{k}, \hat{m}) = (\alpha\theta L r^{-1})^{\frac{\alpha}{1-\alpha}}\hat{m}$. Then, the preferred loan volume under lending rate r is equal to

$$B(\theta, L, m, q, r) = \frac{\alpha + r}{r(1 + r)}\theta L f(\hat{k}, m) - q = \frac{\alpha + r}{r(1 + r)}\theta^{\frac{1}{1-\alpha}} L^{\frac{1}{1-\alpha}} \alpha^{\frac{\alpha}{1-\alpha}} m r^{-\frac{\alpha}{1-\alpha}} - q.$$

In all cases we write $B = B(r)$ for the firm's choice of loan volume, omitting the remaining arguments whenever doing so will not cause confusion.

Remark It is possible to imagine situations in which the optimal level of physical capital is below the already available equity capital q, that is, the firm does not need to finance physical capital by debt. It only has to borrow zm to finance first-period expenditures, that is, to pay for the human capital input. However, it can be shown that limiting lending to *zm* is infeasible as an equilibrium outcome for many important special cases. For instance, under Cobb–Douglas production, banks would be unwilling to lend at a finite rate to such firms. Therefore, I do not consider such cases in this paper. In the numeric examples discussed later, the equilibrium debt levels far exceed current expenditure needs.

The revenue from the loan is $(1 + r)B(r)$ if the realization of S is above S^d (the firms survives) and $SLf(\hat{k}, m) + \hat{k}$ if $S < S^d$. The cost is $(1 + i)B(r)$ in both cases. The expected profit is taken over realizations of S (L is known) and can be written as

$$J^{\mathrm{RB}}(L, r) = \Psi^{-}(S^d)L\hat{f} + X(S^d)\hat{k} + X^{+}(S^d)(1 + r)B(r) - (1 + i)B(r). \quad (9)$$

In Eq. (9), superscript RB refers to relationship banker and $\hat{f}$ is shorthand for the production function value under the optimal choice of the firm. The loan manager chooses r to maximize the right-hand side of Eq. (9) with the knowledge of the loan demand function given by Eq. (8). When this maximization problem has a (finite) solution and, under this solution, the firm equity is priced according to Eq. (3) or (3C), we obtain an equilibrium lending rate for the delegated loan management case for the firms belonging to type (industry) L. This rate is type dependent.

When the wholesale bank sets the rate for all firms itself without delegation, it has the objective function obtained by taking expectation over L of the right-hand side of Eq. (9). That is (superscript AL refers to the arm's-length handling of credit provision),

$$J^{\mathrm{AL}}(r) = \int J^{RB}(L, r)\psi(L)\,\mathrm{d}L, \quad (10)$$

where ψ is the probability density function of the public's (hence also wholesale banks') beliefs about L. In the relationship and wholesale bank cases both, the stock price p (equivalently, the amount of physical capital financed by equity q) of the loan applicant is taken as given.

Functions J^{RB} and J^{AL} both have at most one internal maximum $r^*(q)$ in r for every value of q. It is given by the obvious first-order condition

$$J_r(r^*) = N_r(r^*) - B_r(r^*)i = 0. \quad (11)$$

In Eq. (11), the superscript is dropped for notational economy and N denotes the sum of the first four terms on the right-hand side of Eq. (9) in the relationship banking case, and their Lexpectation as given by the right-hand side of Eq. (10) in the wholesale banking case. Subscripts denote partial derivatives.

To achieve equilibrium, the curves $r^*(q)$ and $q^* = q_0 + p^*(r)$ must intersect in the (q, r) plane. [(Here, p^* is the stock price of the borrower, determined in Sect. 2.2 as a function of lending rate r; this is a function if the price satisfies Eq. (3) and a correspondence if it satisfies Eq. (3C)]. If the curves do not intersect, there is no equilibrium. If they intersect at more than one point, there are multiple equilibria.

2.4.2.1 Definition of equilibrium Every firm sells its shares in the equity market and a debt obligation to a bank. We are looking for a joint equilibrium for the firm share price and the interest rate on the bank loan. This will be a rational expectations equilibrium in which equity market participants condition their investment on lending rate, whereas the parties negotiating a loan condition its size and lending rate on equity price. The loan negotiation is a leader (bank)–follower (firm) subgame embedded in the overall RE equilibrium of simultaneous equity and debt market clearing. The equilibria are naturally split into two categories. In the first, under AL loan management, there is one lending rate for all borrowers. The second is for RB economies, in which there is one lending rate for each borrower type L. Since the bank moves first with its lending rate announcement, inferring firm type based on selected loan volume in the AL case is ruled out, and we also rule out the possibility of such inference from the stock price.[13]

3 Equilibria with and without biased sentiment

3.1 Parameterization of the borrower type space and public beliefs

In the following, we obtain the results in a setup involving only two productivity types, deviating downward or upward from the average (so that $L \in \{L_d, L_u\}$, $L_d < L_u$), in which loan management is either AL or RB for all firms at once. If there were more than two elements in set $\mathbf{L}$, one could also consider different wholesale banks choosing different subsets of $\mathbf{L}$ in which to try out delegation, but this extension is omitted from the present analysis.

Information held by retail investors and wholesale banks alike is parameterized by the value λ, giving the perceived proportion of high-productivity firms in the economy. One example of a situation in which public knowledge of λ is incorrect is when all agents share a biased prior belief.[14] Every firm, although unable to communicate its productivity type credibly to anyone but its relationship banker, is nonetheless able to send an unbiased, even if noisy, public signal about its type. Then, the Bayesian belief update procedure results in a reduction (depending on the relative variances of the signal noise and the prior belief distributions), albeit never complete elimination, of error in the public perception. That is, a portion of the prior bias is preserved even though the signal sent by each firm is unbiased and is processed rationally.

Note that when the solution of Eqs. (3) and (11) is being sought, the relevant value of λ is the one characterizing the beliefs about, and not the actual weight of, the high-productivity industry. This is because the perceived λ enters both the retail

[13] For the latter inference to work, one would need, quite unrealistically, the agents to have complete knowledge of the model and no interfering noise. More generally, it makes no sense to explore the pooling/separating equilibrium issue here, since implementation of a separating equilibrium under the preferences, technologies, and uncertainty distributions dictated by our macro application would be highly artificial in view of associated stability, sensitivity, and robustness problems. Recall also the corresponding remark in Sect. 2.4.

[14] The assumption of common prior beliefs is made to simplify the analysis of *public sentiment* implications. It can be easily relaxed if it is necessary to consider belief differentials across important subcategories of economic agents.

investor and the wholesale banker decision problem (delegated loan managers already know the exact borrower type, so that for them the value of λ is irrelevant). The true λ is important for determining economy-wide aggregates (e.g., investment, bank credit, and average output) after individual decision problems have been solved and equilibrium established.

3.2 Benchmark: equilibrium under unbiased sentiment

We next discuss quantitative properties of equilibria, which were calculated by numerically solving Eqs. (3) and (11) for variables q (equity capital) and r (the lending rate). Recall that the stock price, equal to share capital less the foundation stake ($p = q - q_0$), is in both cases common to all firm types, since retail investors in every stock have the same imperfect information about type as do wholesale banks.

The following functional forms were used throughout the calculations. Retail investors have a negative exponential utility of final wealth with the absolute risk aversion coefficient 0.3. Firms have a Cobb–Douglas production function with physical capital share $\alpha = 1/3$ (the usual value in calibrated macro models). Systemic productivity component S is log-normally distributed with $s = \log S$ having mean -0.125 and standard deviation 0.5. Accordingly, the mean value of S is unity. This choice of the standard deviation is consistent with the following property: in the equilibria obtained for all individual exercises, both high- and low-type producers survive at least for s realizations within two standard deviations from the mean. The foundation stake in every firm is set to 0.2 (which lies between 5 and 7 % of the firm market value in all our exercises). The base cost of funds for the bank, as well as the deposit rate, is 3 %. The values chosen for these last two parameters, as well as other parameters relevant for the corresponding exercise, are reported in the notes to the tables of results.

I begin by showing the results of the equilibrium calculation in the unbiased sentiment case and then discuss the changes caused by either optimistic or pessimistic prejudice.[15]

The results for the unbiased sentiment case are shown in Table 1. As expected, more high-productivity firms (i.e., higher value λ, both perceived and actual as long as there is no prior bias) in the economy means more equity investment, but also higher lending rates (for everyone in the AL case and, on average, for the RB case as well). A less obvious outcome is a decrease in bank credit, investment, and output for each individual type at the same time the aggregate values of these fundamentals grow with λ. This is a sort of "aggregate income effect": the more numerous the high-productivity firms, the less effort needed from each individual producer to attain a given level of aggregate expected output.

Further, looking specifically at the equilibria in the relationship banking environment, lending rates for low-productivity firms fall (moderately) with growing λ, whereas they increase with λ for high-productivity firms. At the same time, higher λ corresponds to higher levels of bank credit, investment, and output in the

[15] All calculations were conducted using Mathematica®.

Table 1 Economic fundamentals in equilibrium with unbiased sentiment

Proportion of high-productivity borrowers:	$\lambda = 0.4$		$\lambda = 0.5$		$\lambda = 0.6$	
Indicator		Aggregate		Aggregate		Aggregate
AL						
q	3.244		3.328		3.398	
r	0.074		0.075		0.076	
B_d	16.391	19.595	16.055	19.823	15.649	19.949
B_u	24.402		23.590		22.815	
k_d	17.256	20.072	17.017	20.324	16.698	20.467
k_u	24.296		23.632		22.980	
y_d	21.088	24.443	20.832	24.809	20.489	25.057
y_u	29.475		28.786		28.102	
RB						
q_d	3.009		3.138		3.267	
q_u	3.389		3.413		3.440	
r_d	0.081		0.080		0.079	
r_u	0.070		0.072		0.075	
B_d	14.304	19.061	14.433	19.516	14.575	19.805
B_u	26.197		24.599		23.292	
k_d	15.055	19.495	15.299	19.988	15.555	20.307
k_u	26.155		24.677		23.475	
y_d	18.718	23.820	18.981	24.446	19.257	24.884
y_u	31.475		29.912		28.635	

The foundation stake q_0 in firm equity is at level 0.2. The cost of lendable funds (deposit rate) is 0.03. For firms of type #, $q_\#$ is total equity capital, $r_\#$ is the borrowing rate, $B_\#$ is the volume of credit taken, $k_\#$ is the total investment in physical capital, $y_\#$ is expected gross output (when the systemic productivity factor takes its expected value of 1), AL is arm's-length loan management, RB is relationship banking (delegated loan management)

low-productivity segment, but lower levels of the same fundamentals in the high-productivity segment.

Note that this effect obtains in a joint equilibrium in equity and debt markets, whereas it would be absent from parallel partial-equilibrium models of the two markets considered separately. In the latter situation, investment and output would always fall with the lending rate (like in the IS equation of the old Keynesian models), and the same is true for the equity price. Looking at Table 1, one sees that our approach renders substantially different reduced-form behavior patterns of the basic fundamentals (cf. the earlier discussion in the penultimate paragraph of the literature review in Sect. 1.2). The loan rate is not just an expression of the price of risk, but also reflects the Stiglitz–Weiss (1981) tradeoff between solvency and revenue from debt service. So, for instance, in the RB case, high net worth producers pay a high interest rate, a phenomenon unknown in standard credit risk theory.

Finally, for each fixed λ, when public sentiment is unbiased, aggregate economic activity is lower in RB economies than in AL ones. This is due to the relatively low common interest rate that the imperfectly informed wholesale banks charge everybody. Although, in RB economies, credit to high-productivity firms is cheaper than in AL ones, so that those firms invest and produce more, credit to low-productivity firms is much more expensive and their output is much lower than in the AL case. Under unbiased sentiment, the latter effect dominates the former. That is, in the world we have created, welfare is not a monotonous function of information quality: on average, the involvement of loan managers who know the borrower type generates less output than their absence. So, in our model world, which is not unlike the developed economies shortly before the outbreak of the latest crisis, banks may be tempted to refrain from the costly use of intermediary agents with superior information and instead grant loans based on general formal rules (this is the essence of the AL approach).[16]

Although formally supported by a different type of model, this outcome is reminiscent of situations described in the classical financial intermediation literature dealing with how information quality about borrowers affects the social optimality of investment choices (see, e.g., Sharpe 1990; Rajan 1992). Consequently, my environment can accommodate adverse selection and moral hazard effects. A deeper exploration of the welfare role played by the lending regime would require a richer macro model.

3.3 Extension: bias in public perception and equity bubbles

The results illustrating the corresponding economic sentiment effect are set out in Table 2. Within each borrower type, determination of the equilibrium equity price and lending rate depends on the perception of (i.e., not the actual) λ. The difference between subjective beliefs and reality matters for the observed economic aggregates. As expected, aggregate bank credit, as well as investment and output, grow along with the actual proportion of high-productivity firms. On the other hand, however, for every fixed value of actual λ, economic activity *falls with growing perceived λ*. This outcome is intuitive and can be roughly explained as follows. When banks believe that there are more high-productivity borrowers, they also feel justified in charging higher lending rates. Since only high performers choose high investment volumes when credit is expensive, the latter depresses aggregate investment to a greater extent than everyone expects.

In other words, incorrect economic sentiment incurs an aggregate cost. In this respect, RB economies are slightly less sensitive to prior bias than are AL economies, and it may also occasionally happen that the RB output under a particular sentiment value exceeds the AL output (as when perceived λ is 0.4 and the actual one is 0.6 in our example). In all cases, Table 2 suggests that, for a fixed absolute size of sentiment error, it is socially preferable for people to be pessimistic. This follows from comparing

[16] Interestingly, the outcome will be reversed under some realizations of prior prejudice. This should act as a warning that in a general equilibrium environment, asset bubbles caused by cognitive aberration have the potential to reverse conventional findings of adverse selection/moral hazard microeconomics.

Table 2 Main fundamentals under changing sentiment

Perceived proportion of high-productivity borrowers:	$\lambda = 0.4$			$\lambda = 0.5$			$\lambda = 0.6$		
True value of λ:	0.4	0.5	0.6	0.4	0.5	0.6	0.4	0.5	0.6
Indicator									
AL									
q	3.244	3.244	3.244	3.328	3.328	3.328	3.398	3.398	3.398
r	0.074	0.074	0.074	0.075	0.075	0.075	0.076	0.076	0.076
B	19.595	20.396	21.197	19.069	19.823	20.576	18.516	19.232	19.949
k	20.072	20.776	21.480	19.663	20.324	20.986	19.211	19.839	20.467
y	24.443	25.282	26.120	24.014	24.809	25.605	23.534	24.296	25.057
RB									
q_d	3.009	3.009	3.009	3.138	3.138	3.138	3.267	3.267	3.267
q_u	3.389	3.389	3.389	3.413	3.413	3.413	3.440	3.440	3.440
r_d	0.081	0.081	0.081	0.080	0.080	0.080	0.079	0.079	0.079
r_u	0.070	0.070	0.070	0.072	0.072	0.072	0.075	0.075	0.075
B	19.061	20.396	21.197	18.499	19.516	20.533	18.062	18.933	19.805
k	19.495	20.605	21.715	19.050	19.988	20.926	18.723	19.515	20.307
y	23.820	25.096	26.372	23.353	24.446	25.539	23.009	23.946	24.884

The foundation stake q_0 in firm equity is at level 0.2. The cost of lendable funds (deposit rate) is 0.03. For firms of type #, $q_\#$ is total equity capital, $r_\#$ is the borrowing rate; variables without subscripts denote economy-wide aggregates; B is the volume of credit taken, k is the total investment in physical capital, y is expected gross output (when the systemic productivity factor takes its expected value of 1), AL is arm's-length loan management, RB is relationship banking (delegated loan management)

economic activity for, say, the combination actual $\lambda = 0.4$, perceived $\lambda = 0.5$ with the combination actual $\lambda = 0.5$, perceived $\lambda = 0.4$, and so forth.

It remains to be seen to what extent this particular result, and others, are influenced by the employed orthodox efficient market paradigm of equity pricing. Specifically, this paradigm often dictates a strong mutual reinforcement of interest rate and equity reactions to exogenous shocks. Moreover, the observed sensitivity of economic activity values to sentiment changes is even higher. For instance, under a 0.1-size change of sentiment (i.e., the perceived λ value), the interest rate also changes by roughly 0.1 percent, but the output values shift by 3 % or more.

In the next section, we test the ability of the constructed model to address the real effects of macroprudential regulation of financial intermediaries.

4 Macroprudential capital charges, borrower liability, and economic activity

4.1 Systemic risk and macroprudential response

In the present model, systemic risk is embodied by the aggregate random productivity component S of TFP variable A. Note that from the viewpoint of a given lender or a given borrower, it makes no difference whatsoever whether, at one extreme, S is common for all producers, or at the other, there is an independent identically distributed individual risk factor for each firm, or, intermediately, there is some imperfect correlation between those identically distributed risks. Indeed, a firm manager cares only about the firm's own production, taking the physical capital level and the interest rate set by the bank as given. A relationship banker in an RB economy is a mirror image of the firm manager, caring only about the performance of himself, and not the economy-wide aggregates. Finally, in an AL economy, the lending rate is set by the wholesale bank, which uses a preset rule for this purpose that employs a demand curve with a distributed parameter—borrower type L, possibly distorted by sentiment—but without comparing different individual borrowers.

Consequently, in the present setting, correlations between Srealizations of individual producers are a concern only for the policymaker. Systemic risk is highest when the correlation is perfect, that is, there is, essentially, a common random disturbance S affecting all firms at once.[17] For simplicity, this is the only case for which we will derive numerical results. Informally (with quantitative policy loss functions relegated to future research), we assume that the policymaker cares about three factors: economic activity, probability of default, and LGD. For a given equilibrium, the latter two factors become straightforward functions of the distribution of systemic risk factor S. These will also be the factors intended to be mitigated by means of a macroprudential policy instrument.

There is a growing consensus in financial crisis research that debt volume grows and its quality deteriorates much faster in the runup to a financial crisis than in normal times. Therefore, policymakers have for some time been looking for an adequate means to curtail unusual debt expansions and prevent credit bubbles without choking "genuine" growth. At the moment, reliable strategies for separating bubbles from sustainable growth are unavailable. So, most probably, in the pursuit of financial stability, most regulators would resort to simple penalties for suspicious credit expansions by mandating additional capital cushions against lending that visibly exceeds the accepted target. That is, the most likely macroprudential policy instruments are capital requirements on—and, hence, additional costs of—incompletely collateralized loans, requirements that would grow faster than proportionally with loan volume. In the present model, one can accommodate such an instrument by replacing the linear cost-of-funds term in the bank objective function with a linear-quadratic term that contains a surcharge on the loan volume in proportion to growing borrower leverage.

Formally, let us introduce the following macroprudential control mechanism into the model. If the target level of physical capital of the borrower is k and the loan size is B, the bank is subject to an additional charge (in the form of regulatory capital) that leads to extra funding costs equal to $\frac{a}{2k}B^2$, where a is a positive constant. That is, the

[17] Observe that if, in comparison to the model here that encompasses only two representative firms—one of each type—we had a large number of small-size borrowers, perfectly independent firm-level risks would mean no systemic risk whatsoever. In our two-firm setting, the exact size of systemic risk depends, beside correlation, on the relative size of two representative borrowers. This issue is not explored any further in the present paper.

funding cost term $(1+i)B$ in Eq. (9) [and Eq. (10)] is replaced with $(1+i)B + \frac{a}{2k}B^2$. This means that

(a) funding costs are growing as a convex function of the loan volume and not linearly as in the original Eq. (9);
(b) the surcharge is proportional to the product of the loan volume and the borrower's debt-to-physical capital ratio; every additional unit of credit is penalized unless offset by physical capital collateral financed with equity;
(c) the unit of penalty for uncollateralized credit is a, usually a single-digit percentage point amount.[18]

Our formal representation of the macroprudential policy is similar in spirit to the within-period snapshot rule presented in Angelini et al. (2010) (the dynamic element of the latter would be useless in our two-period setup). Under this type of macroprudential policy, the loan supply becomes more elastic. In the RB case, the bank will charge each borrower type an interest rate based on a specific supply curve. Only in the "corner" AL case, in which the bank does not distinguish between types ex ante, will rates increase uniformly. Still, equilibrium loan volumes will be different for different lender types in all cases. (Recall, nevertheless, that we view type identification by volume alone as unrealistic; cf. Sect. 2.4 and the beginning of Sect. 3, including note 13.)

Remark: comparison with conventional capital requirement policies (Basel II) It is natural to inquire how the capital surcharge rule introduced above differs from the capital requirement practice of the recent precrisis past. Obviously, the prime conceptual difference is that Basel I and II rules were microprudential, that is, they had not been designed to pay any special attention to systemic risk.[19] Consequently, these rules are not of immediate concern here. However, in terms of the present model with just two representative borrowers, can one find at least formal overlaps, for example, with the Basel II mechanism? The cornerstone of capital requirements according to Basel II is the level of risk-weighted assets, of which, for our purposes, the key part would be the risk-adjusted sum of exposures across the loan book. The regulatory capital formula uses the tail size of the loss distribution constructed for that same collection of assets. In other words, the key ingredient of Basel II is the second moment of the imputed loss distribution. In our setting, this would be the corresponding summary statistic of the systemic risk factor S (or total productivity risk $A = SL$ if the bank does not delegate to loan officers who know L). On the contrary, our macroprudential tool is "nonparametric," that is, defined without a direct reference to the distributional parameters of risk factors. Instead, by employing the borrower credit-to-own capital variable as a driving factor of additional loan funding cost, we choose an instrument based only on the hard evidence of the balance sheet, particularly indicators of leverage. That is, unless one sets risk weights directly based on borrower leverage, our

[18] In our calculations, we use $a = 0.01$. This corresponds to adding 1 % to funding costs of a loan of whose face value 50 % is secured by the borrower's physical capital.

[19] As we know, the most notorious macro failure of the Basel rules is their tendency to support the pro-cyclicality of capital requirements instead of suppressing it. Another, less conspicuous, but equally fundamental weakness in the rules from the systemic risk perspective is the lack of separation between the precautionary buffer and the leverage-brake functions of capital requirements.

macroprudential instrument is complementary to Basel II: loan funding costs for the bank are not a function of a borrower's return volatility, but of the borrower's "worst-case" recovery rate. In any event, our rule is in line with the loan-to-value restrictions currently considered the central element of macroprudential regulation policies.

4.2 Macroprudential policy effects in a no-bubble environment

The calculation of equilibrium under prudential capital surcharges can proceed in the same way as before, with only term N in Eq. (11) needing to be modified so as to reflect the new definition of funding costs. We show the results for the base case of equal borrower type weights and no prior bias in Table 3. For the sake of accurate comparison with the original model without prudential policy instruments, we assume that the extra funding costs carried by the banks are turned back over to the private sector in the form of transfers (e.g., tax relief) and, therefore, are included in the aggregate output measures.

Table 3 shows that macroprudential policies in the defined form achieve one immediate goal: reduction of the LGD level. On the downside, they are a significant extra component in the price of credit and a heavy burden on economic activity. The bulk of this burden is carried by high-productivity borrowers, so that their distance from low-productivity ones in terms of investment and output is now smaller. In the present model, high-productivity firms take on more risk and default more frequently than those in the low-productivity segment. So, if the objective of macroprudential policies is to check the expansion of the riskiest segments of the bank loan market, it is being achieved through dampening economic activity in the high-productivity segment. On the other hand, if the objective of the macroprudential instrument involves stabilizing or reducing the number of defaults (this can be the case if defaults carry a negative externality that enters the social planner's objective function), then the instrument is clearly counterproductive: the number of defaults is now higher. Since the relationship banking regime in general is more favorable to high-productivity firms, the costs of new policies are higher in RB economies as well. The presence of biased sentiment (not shown) does not much change the results.

To sum up, the main advantage of the considered capital surcharge is that, by achieving a reduction of imprudent leverage by highly productive borrowers, it helps bring down average LGD figures (conditioned on default by the more productive segment; the default probability of the less productive segment is too low to matter in any case).

On the contrary, the main problems associated with the use of the instrument can be identified as

- a uniform and significant increase in lending rates for all borrowers;
- an increase in default rates (which may be a problem if such are associated with welfare externalities of concern to the policymaker); and
- high sensitivity of investment and output to small changes of the capital charge rate (this has to do with the additional transmission channel through equity markets).

Naturally, the actual *raison d'être* for a macroprudential tool of the above type is its ability to stabilize inflation and output in the medium run, that is, it can manifest fully

Table 3 Economic fundamentals in equilibrium with and without prudential capital surcharges

Prudential capital surcharge:	No		Yes	
Indicator		Aggregate		Aggregate
AL				
q	3.328		2.999	
r	0.075		0.085	
Default probability				
L_d	+0		+0	
L_u	0.025		0.039	
Loss given default	20.593		17.193	
B_d	16.055	19.823	13.304	16.637
B_u	23.590		19.970	
k_d	17.017	20.324	14.101	16.984
k_u	23.632		19.867	
y_d	20.832	24.809	17.684	21.201
y_u	28.786		24.709	
RB				
q_d	3.138		2.756	
q_u	3.413		3.099	
r_d	0.080		0.094	
r_u	0.072		0.081	
Default probability				
L_d	+0		+0	
L_u	0.022		0.034	
Loss given default	22.937		19.586	
B_d	14.433	19.516	11.437	16.230
B_u	24.599		21.022	
k_d	15.299	19.988	12.116	16.540
k_u	24.677		20.964	
y_d	18.981	24.446	15.528	20.717
y_u	29.912		25.906	

Results are shown for the perceived share $\lambda = 0.5$ of high-productivity borrowers and no prior bias. The base capital surcharge $a = 0.01$ corresponds to 1 percent per first unit of credit uncollateralized by physical capital. The foundation stake q_0 in firm equity is at level 0.2. The cost of lendable funds (deposit rate) is 0.03. For firms of type #, $q_\#$ is total equity capital, $r_\#$ is the borrowing rate, $B_\#$ is the volume of credit taken, $k_\#$ is the total investment in physical capital, $y_\#$ is expected gross output (when the systemic productivity factor takes its expected value of 1), AL is arm's-length loan management, RB is relationship banking (delegated loan management). LGD is the expectation with respect to productivity distribution conditioned on default

only in a dynamic environment. Nevertheless, the consequences of its application in terms of expensive credit and low investment are likely to carry over from our present two-period to a multi-period model. Therefore, a capital charge mechanism that could minimize the side effects on quality borrowers would be always welfare improving. To investigate this possibility, we considered a variant of the present model with proportional liability rules for firm management remuneration, which effectively induces unlimited-liability decisions about capital structure and input purchases. The essential property desired from this alternative incentive scheme is for the manager to discriminate between different earning realizations not just in good times (survival) but in bad times (default) as well.

Our conjecture is that the road toward a welfare-improving capital requirement policy goes in the direction of encouraging lending to personally liable borrowers and penalizing excessive exposure to borrowers with conventional limited liability. Supporting evidence in the present setting is provided by a comparison of the outcomes of the benchmark model (limited-liability borrowers) with those of the model under the condition of unlimited liability behavior (cf. the remark at the end of Sect. 2.2).[20]

The comparison (under equal productivity type weights and unbiased public sentiment) is summarized in Table 4, which shows that inducing unlimited liability behavior has four major consequences compared to the benchmark:

(a) the lending rates of both productivity types are quite similar in the RB case and approach the common lending rate of the AL case;
(b) there is a sharp increase in the equity value of the high-productivity type, and a minor decrease in the equity value of the low-productivity type;
(c) the default probability of the high-productivity type falls substantially, whereas for the low-productivity type, although formally increasing, this probability remains negligible; and
(d) there is a minor reduction of economic activity compared to the level in the presence of limited-liability borrower behavior.

If the policymaker's prime concern is to find a macroprudential policy without a major negative impact on economic activity, unlimited-liability behavior by borrowers, if such could be achieved, would have an advantage over the previously considered convex capital surcharge instrument, provided that it was possible at the same time to encourage delegated loan management in banks. Indeed, suppose that unlimited-liability behavior is impossible to implement in the AL regime, but possible in the RB regime (i.e., the relationship banker is able to influence the manager incentive structure in the borrowing firm). In that case, the tradeoff for the macroprudential policy is between a convex capital surcharge in wholesale banks against unlimited borrower liability behavior in the RB regime without capital surcharges. Comparing the upper-right panel of Table 3 with the lower-right panel of Table 4, one sees that the losses in economic activity caused by abandoning limited liability are more than compensated for by the possibility of avoiding additional capital requirements, with the extra bonus of reducing default rates.

[20] Recall that the considered behavior of the borrower firm does not mean that the lender receives full repayment in all states of nature, meaning that in adverse states of nature (output less than debt service, i.e., default), a part of the compensation comes from the borrower's private wealth. As before, in default the jointly available assets of the firm and its management are insufficient to service the debt. All that is assumed here is that the firm manager compensation is an affine function of firm earnings less debt service. In such a case, the manager would select production inputs *as if* the firm operated under unlimited liability.

Table 4 Economic fundamentals in equilibrium when borrower incentives replicate either limited or unlimited liability conditions

Borrower incentives:	LL-consistent		UL-consistent	
Indicator		Aggregate		Aggregate
AL				
q	3.328		3.916	
r	0.075		0.075	
Default probability				
L_d	+0		+0	
L_u	0.025		0.00086	
B_d	16.055	19.823	15.505	18.916
B_u	23.590		22.326	
k_d	17.017	20.324	17.053	20.047
k_u	23.632		23.042	
y_d	20.832	24.809	20.870	24.535
y_u	28.786		28.200	
RB				
q_d	3.138		3.072	
q_u	3.413		3.881	
r_d	0.080		0.076	
r_u	0.072		0.077	
Default probability				
L_d	+0		0.000051	
L_u	0.022		0.00062	
B_d	14.433	19.516	15.879	18.522
B_u	24.599		21.164	
k_d	15.299	19.988	16.606	19.256
k_u	24.677		21.907	
y_d	18.981	24.446	20.390	23.684
y_u	29.912		26.979	

Results are shown for the perceived share $\lambda = 0.5$ of high-productivity borrowers and no prior bias. The foundation stake q_0 in firm equity is at level 0.2. The cost of lendable funds (deposit rate) is 0.03. For firms of type #, $q_{\#}$ is total equity capital, $r_{\#}$ is the borrowing rate, $B_{\#}$ is the volume of credit taken, $k_{\#}$ is the total investment in physical capital, $y_{\#}$ is expected gross output (when the systemic productivity factor takes its expected value of 1), AL is arm's-length loan management, RB is relationship banking (delegated loan management). Borrower management incentives are either consistent with limited liability (LL) or imitate unlimited liability (UL)

4.3 Economic activity under biased sentiment and macroprudential policy reaction

We now discuss macroprudential intervention outcomes in the presence of a bubble. The experiment involves a positive bubble generated by a prior belief of 40 % low-productivity producers, whereas their true proportion is 60 %. The results are shown in Table 5.

Table 5 Bubble, bank business model, and regulatory stance

Bubble: Y/N	Equity price		Output		Interest rate		Default probability			
	Yes	No	Yes	No	Yes	No	Yes		No	
	Productivity type						Low	High	Low	High
AL										
Baseline	3.198	3.044	23.534	24.443	7.57 %	7.40 %	+0	0.0163	+0	0.0384
Relative difference between bubble and no-bubble outcomes (%)	5.07		−3.72							
Macorprudential instrument	2.871	2.731	20.170	20.961	8.54 %	8.37 %	+0	0.0263	+0	0.0580
Relative difference between bubble and no-bubble outcomes (%)	5.13		−3.77							
Downside borrower risk	3.724	3.722	23.606	24.038	7.51 %	7.41 %	+0	0.0007	+0	0.0009
Relative difference between bubble and no-bubble outcomes (%)	0.04		−1.80							
RB										
Baseline	3.136	2.961	23.009	23.820	7.69 %	7.51 %	+0	0.0153	+0	0.0319
Relative difference between bubble and no-bubble outcomes (%)	5.92		−3.41							
Macorprudential instrument	2.776	2.631	19.430	20.234	8.77 %	8.53 %	+0	0.0241	0.0001	0.0471
Relative difference between bubble and no-bubble outcomes (%)	5.53		−3.97							
Downside borrower risk	3.337	3.052	22.524	22.714	7.58 %	7.73 %	+0	0.0005	0.0008	0.0007
Relative difference between bubble and no-bubble outcomes (%)	9.36		−0.84							

Al arm's-length loan management, *RB* relationship banking (delegated loan management)

In all three lending regimes considered (baseline with limited borrower liability and without macroprudential capital charges, under macroprudential capital charges, and under proportional downside risk of the borrower management), the bubble means higher equity prices compared to the no-bubble benchmark. The extent of equity overvaluation is somewhat higher under delegated loan management. Average output, on the other hand, is lower under the bubble. A little surprisingly, default probabilities are lower compared to the benchmark. This reduction is caused by lower interest rates charged to high-productivity firms under optimistic sentiment.

Under the arm's-length bank business model, we observe that in the presence of the macroprudential instrument, bubble removal (restoration of unbiased sentiment) has a marginally stronger impact on both the equity price and output than in the baseline case. In the case of a positive bubble, the equity price falls and output rises when the bubble disappears, the same as in the baseline case. The difference between bubble and no-bubble interest rates is also mildly stronger. Unfortunately, this effect is achieved at the cost of reduced economic activity.

Under loan management delegation (relationship banking), the quantitative effects of the macroprudential capital charges are even less impressive, since, compared to the baseline, both the output gain and the bubble price correction are smaller. The adverse absolute impact on output is even more severe than in the AL business model. Apparently, the results of macroprudential policies enacted in a relationship banking environment can be very disappointing due to their inability to exploit the informational advantages of delegated loan managers.

On the other hand, enforcing firm manager downside risk, if feasible, could both reduce the sensitivity of the real economy to the asset price bubble and (the same as in the no-bubble benchmark) improve absolute output levels.

4.4 Bubble containment alternatives: macroprudential versus monetary policy

Having seen the consequences of bubble containment with the introduced macroprudential instrument, we now compare those consequences with the results of a more traditional monetary policy tool. We model monetary policy as an exogenous simultaneous change in the funding cost and the deposit interest rate value i. That is, while monetary policy affects the linear component of banks' financing costs (irrespective of whether this financing comes from deposits or the money market), the macroprudential tool leaves this linear component unchanged but adjusts the curvature of the nonlinear cost component.

For concreteness, calculations were conducted for our basic limited borrower liability case and unbiased sentiment. Variations in the value of i resulted in equilibria that, with respect to variables characterizing economic activity, were hard to distinguish from the ones obtained from the baseline case by introducing the macroprudential tool of the preceding subsections. In particular, under the used parameterization, we found that economic activity returns, more or less, to the baseline level when the macroprudential policy tested in our simulations is compensated for by a 0.5 % monetary policy easing (reduction of i). Equivalently, average output, investment, equity prices, and

Table 6 Macroprudential capital surcharges versus monetary policy tightening

Policy stance:	Macroprudential instrument		Interest rate hike	
Indicator		Aggregate		Aggregate
AL				
q	2.999		3.023	
r	0.085		0.085	
Default probability				
L_d	+0		+0	
L_u	0.039		0.035	
Loss given default	17.193		17.083	
B_d	13.304	16.637	13.285	16.574
B_u	19.970		19.862	
k_d	14.101	16.984	14.106	16.950
k_u	19.867		19.795	
y_d	17.684	21.201	17.690	21.163
y_u	24.709		24.637	
Bank earnings	0.899965		0.817987	
RB				
q_d	2.756		2.791	
q_u	3.099		3.119	
r_d	0.094		0.093	
r_u	0.081		0.082	
Default probability				
L_d	+0		+0	
L_u	0.034		0.031	
Loss given default	19.586		18.013	
B_d	11.437	16.230	11.534	16.195
B_u	21.022		20.855	
k_d	12.116	16.540	12.241	16.536
k_u	20.964		20.831	
y_d	15.528	20.717	15.659	20.711
y_u	25.906		25.764	
Bank earnings	0.8949373		0.816335	

Results are shown for the perceived share $\lambda = 0.5$ of high-productivity borrowers and no prior bias. The base capital surcharge $a = 0.01$ corresponds to 1 % per first unit of credit uncollateralized by physical capital. The foundation stake q_0 in firm equity is at level 0.2. The cost of lendable funds (and the deposit rate) is 3 % under macroprudential measures and 3.5 % in their absence (the "Interest rate hike" columns). For firms of type #, $q_{\#}$ is total equity capital, $r_{\#}$ is the borrowing rate, $B_{\#}$ is the volume of credit taken, $k_{\#}$ is the total investment in physical capital, $y_{\#}$ is expected gross output (when the systemic productivity factor takes its expected value of 1), AL is arm's-length loan management, RB is relationship banking (delegated loan management). Borrowers have limited liability

lending rates all react to the macroprudential tool approximately the same way as they would to a 0.5 % increase in the value of interest rate i. The results of the latter double experiment are shown in Table 6.

What is the difference, then, between the two policy effects? Inspection of Table 6 suggests that there may be a minor increase in default rates and LGD (in the high-productivity, hence risky, borrower segment) under macroprudential tightening compared to the monetary policy one, but, given the chosen level of precision, this increase is barely outside confidence bounds. Much more important is the comparison of effects on bank earnings: they are significantly higher under the macroprudential tightening than in the key rate-hike policy case. This indicates that, should the monetary authority decide to engage in countercyclical action, banks would prefer macroprudential activism to standard monetary policy. This is because monetary policy affects bank funding costs uniformly and jointly with the firm equity and household deposit markets. On the other hand, the macroprudential tool immediately affects only the lender–borrower link and allows the bank to pass higher funding costs on to the real sector more easily. This is especially pronounced in the relationship banking case because the price discrimination opportunity offered to the loan manager based on his superior knowledge of the borrower allows the bank to extract higher rents.

Thus, both policy instruments—macroprudential and monetary—contribute superficially to suppressing equity price bubbles, but are disappointing in regard to their effects on other credit excess phenomena such as default probabilities. Quit importantly, their performance on the real economy side of bubble containment is marred by the overall dampening impact on investment and output.

5 Conclusion

I constructed a model of the financial sector that is an interface between conventional optimizing general equilibrium macro models and partial equilibrium models of financial intermediation driven by information asymmetries. The presence of a macroprudential instrument in this model reduces equilibrium fragility and fulfills the formal task of suppressing LGD levels and excessive equity price volatility, but fails to reduce default rates. Most importantly, it entails a uniform and significant increase in lending rates for all borrowers, resulting in a tangible output loss. The effect obtains because the most risky borrower segment, that is, the one targeted by the policy, is also the one with the highest ex ante performance. Macroprudential capital surcharges mean that those borrowers experience a disproportional, as opposed to the low-productivity segment, increase in the price of credit. This is a consequence of an additional transmission channel through physical capital markets: after an initial increase in the lending rate, banks have to raise the rate even higher because physical capital as collateral has become cheaper. This mechanism is at work regardless of either the sign or the size of a possible equity bubble.

Both the strength and the robustness of my results with respect to the size of bubble in the physical capital market are due to the mutually reinforcing reactions to shocks by both equity and debt markets. Irrespective of bubbles, the economy demonstrates a high degree of sensitivity of investment and output to small changes in the capital surcharge rate.

A monetary policy tool aimed at the same equity bubble, although it mostly impacts the bank profit side and has very little effect on the producer/borrower side, is slightly better at reducing default frequency, but slightly worse than the macroprudential tool when it comes to lowering average LGD.

As a further observation, and one worthy of more research, is that the convex macroprudential capital charge on bank loans helps the investors coordinate on an equilibrium mix of equity and debt financing in situations in which equilibria do not exist in the absence of this instrument. In my model, this is particularly likely to occur when firm productivity types are distributed very unevenly or when public economic sentiment is highly biased.

Since macroprudential regulation in its current form affects only lenders, it has no direct influence on the behavior of borrowers caused by limited liability. However, adverse aggregate consequences of investment decisions by producers who are indifferent between varying degrees of insolvency (do not distinguish between "dead" and "deader" states of the firm that is unable to repay the loan) are more pronounced than the consequences due to the absence of a macroprudential response to a bubble. Accordingly, policies able to replicate downside risk on the borrower side are likely to be the next fundamental challenge for financial regulation. The resulting equilibria in a modified "proportional liability" regime also entail reduced default rates for the most risky borrowers, compared to the pure limited liability case. More generally, there seem to be limits, in terms of economic activity and ex ante welfare costs, to promoting financial stability through policies directed at *credit providers*. At the same time, policies with the same ultimate objective of systemic risk containment, but directed at *credit consumers*, continue to be largely unexplored (let alone exploited). My results indicate that the potential benefit gained by reorienting from regulating credit supply to educating credit demand may be worthwhile, notwithstanding numerous implementation difficulties.

Acknowledgments I thank seminar participants at the Czech National Bank and the European Central Bank MARS network, participants at the 2011 Swiss Society for Financial Market Research Annual Conference, and an anonymous referee for helpful comments.

References

Angelini, P., Neri, S., Panetta, F.: Grafting Macroprudential Policies in a Macroeconomic Framework. Choice of Optimal Instruments and Interaction with Monetary Policy. Mimeo, Bank of Italy (2010)

Aoki, K., Proudman, J., Vlieghe, G.: House prices, consumption, and monetary policy: a financial accelerator approach. J Financ Intermed **13**(5), 414–435 (2004)

Bernanke, B.S., Gertler, M., Gilchrist, S.: The financial accelerator in a quantitative business cycle framework. In: Taylor, J.B., Woodford, M. (eds.) Handbook of Macroeconomics, 1st edn, chap. 21, vol. 1, pp. 1341–1393. Elsevier, Amsterdam (1999)

Besanko, D., Thakor, A.V.: Competitive equilibrium in the credit market under asymmetric information. J Econ Theory **42**, 167–182 (1987)

Blum, J., Hellwig, M.: The macroeconomic implications of capital adequacy requirements for banks. Eur Econ Rev **39**(3–4), 739–749 (1995)

Bonaccorsi di Patti, E., Dell'Ariccia, G.: Bank competition and firm creation. J Money Credit Bank **36**(2), 225–251 (2004)

Borio, C.: Towards a Macroprudential Framework for Financial Supervision and Regulation? BIS WP No. 128 (February) (2003)

Christiano, L.J., Motto, R., Rostagno, M.: Shocks, structures or policies? The euro area and the US after 2001. J Econ Dyn Control **32**(13), 2476–2506 (2008)

Covas, F., Fujita, S.: Procyclicality of capital requirements in a general equilibrium model of liquidity dependence. Int J Cent Bank **6**(5), 137–173 (2010)
Dell'Ariccia, G., Marquez, R.: Information and bank credit allocation. J Financ Econ **72**, 185–214 (2004)
de Walque, G., Pierrard, O., Rouabah, A.: Financial (in)stability, supervision and liquidity injections: a dynamic general equilibrium approach. Econ. J. **120**, 1234–1261 (2010)
Diamond, D.: Financial intermediation and delegated monitoring. R Econ Stud **51**(3), 393–414 (1984)
Diamond, D., Rajan, R.: A theory of bank capital. J Finance **55**(6), 2431–2465 (2000)
Donaldson, J.B., Gershun, N., Giannoni, M.P.: Some Unpleasant General Equilibrium Implications of Executive Incentive Compensation Contracts. NBER Working Paper Series, vol. w15165 (July) (2009)
He, Z., Krishnamurthy, A.: Intermediary Asset Pricing, Working Paper. Chicago Booth and Northwestern University (2010)
Hui, O.-Y.: Optimal contracts in a continuous-time delegated portfolio management problem. Rev Financ Stud **16**(1), 173–208 (2003)
Iacoviello, M.: House prices, borrowing constraints, and monetary policy in the business cycle. Am Econ Rev **95**(3), 739–764 (2005)
International Monetary Fund (IMF): Lessons for monetary policy from asset price fluctuations. In: World Economic Outlook, Chapter **3** (October) (2009)
Kraft, H., Korn, R.: Continuous-time delegated portfolio management with homogeneous expectations: can an agency conflict be avoided? Financ Mark Portf Manag **22**(1), 67–90 (2008)
Mendoza, E.: Sudden stops, financial crises, and leverage. Am Econ Rev **100**, 1941–1966 (2010)
Monacelli, T.: New Keynesian models, durable goods, and collateral. J Monet Econ **56**, 242–254 (2009)
Rajan, R.: Insiders and outsiders: the choice between informed and arm's-length debt. J Finance **47**, 1367–1400 (1992)
Repullo, R., Suarez, J.: The Procyclical Effects of Bank Capital Regulation, WP 2010–29S. Tilburg University, Center for Economic Research (2010)
Santomero, A.: Modeling the banking firm: a survey. J Money Credit Bank **16**(5), 576–602 (1984)
Saunders, A., Walter, I.: Financial architecture, systemic risk, and universal banking. Financ Mark Portf Manag **26**(1), 39–59 (2012)
Sharpe, S.: Asymmetric information, bank lending, and implicit contracts: a theoretical model of customer relationship. J Finance **45**, 1069–1087 (1990)
Stiglitz, J., Weiss, A.: Credit rationing with imperfect information. Am Econ Rev **71**(3), 393–410 (1981)
Townsend, R.M.: Optimal contracts and competitive markets with costly state verification. J Econ Theory **22**, 265–293 (1979)

Author Biography

Alexis Derviz is a Principal Economist in charge of a team of international economy analysts at the Czech National Bank. He received a Ph.D. in mathematical physics from the Leningrad (St. Petersburg) State University in Russia. He later studied economics at Charles University in Prague, Czech Republic, and taught industrial organization and international finance at the graduate school of economics of that university. He then joined the Czech National Bank to lead a research team on international monetary issues. He contributed to and co-edited several books on economies in transition, later authored a number of refereed journal articles on asset pricing, exchange rates and international banking.