

A General Equilibrium Theory of Banks' Capital Structure*

Douglas Gale[†]

New York University

Piero Gottardi[‡]

University of Essex, University of Venice and CEPR

Revised December 26, 2019

Abstract

We develop a general equilibrium theory of the capital structures of banks and firms. The liquidity services of bank deposits make deposits a “cheaper” source of funding than equity. In equilibrium, banks pass on part of this funding advantage in the form of lower interest rates to firms that borrow from them. Firms and banks choose their capital structures to balance the benefits of debt financing against the risk of costly default. An increase in the equity of a firm makes its debt less risky and that in turn reduces the risk of the banks who lend to the firm. Hence there is some substitutability between firm and bank equity. We find that firms have a comparative advantage in providing a buffer against systemic shocks, whereas banks have a comparative advantage in providing a buffer against idiosyncratic shocks.

*We thank Franklin Allen, Cyril Monnet, the Editor, Xavier Vives, and three anonymous referees for very helpful comments and suggestions. We also thank the participants at seminars and conferences at New York University, the University of Cambridge, the University of Essex, the Stockholm School of Economics, UCLA, IMPA, Galatina and the Copenhagen Business School for their comments and questions. We are especially grateful to Sukjoon Lee for correcting errors in the proofs of Propositions 8 and 9. The second author acknowledges financial support from MUR.

[†]Department of Economics, New York University, 19 West 4th Street, New York, NY 10012 USA. E-mail: douglas.gale@nyu.edu

[‡]Department of Economics, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom. E-mail: piero.gottardi@essex.ac.uk

1 Introduction

Bank capital plays a central role in assuring the safety and soundness of the banking system. Recent work has begun to recognize the general equilibrium dimension of the capital structure decision. Firms and banks choose their capital structures independently, but the firms' decisions have implications for both firms and banks. The riskiness of a bank loan depends on the capital structure of the firm to which it is made. Risk originates in the corporate sector and flows down to the banking sector. Firm equity acts as a buffer that absorbs shocks and prevents costly financial distress. A highly leveraged firm is more likely to experience financial distress, other things being equal, than a firm with low leverage. The more equity there is in a firm's capital structure, the "safer" the bank lending to that firm will be. Hence, the "safety and soundness" of banks depends on the capital structures of their borrowers, as well as on their own capital structures. In this sense, the equity component of a firm's capital structure does double duty: it protects the firm against default and it shields the lending bank from default. This general effect—the "supply chain" of finance, as Gornall and Strebulaev (2017) call it—explains why the efficient allocation of equity capital between the banking and corporate sectors is a general equilibrium problem.

Our main objective is to investigate the efficient allocation of equity capital between the banking and corporate sectors in a competitive general equilibrium setting. We study an economy consisting of consumers, firms and banks. Consumers are the initial owners of capital goods and the ultimate buyers of consumption goods. Firms have access to a variety of risky technologies, which use capital equipment to produce consumption goods. Firms issue equity to households and borrow from banks in order to fund the purchase of capital equipment and choose which technology to operate. Banks issue equity and deposits to households to fund loans to firms and choose the composition of their loan portfolio. Households purchase equity and make deposits in banks to fund their future consumption. Only households can invest in equity.

Firms and banks are restricted in the securities they can issue. Firms are restricted to issuing debt and equity. Banks are restricted to taking deposits and issuing equity. In this sense, markets are incomplete. Nonetheless, the set of potential securities that can be issued is large because the risk characteristics of these securities depend on the decisions taken by firms and banks. If firms choose different technologies or capital structures, they change the risk characteristics of their debt and equity. Similarly, if banks choose different loan portfolios or capital structures, they change the risk characteristics of their deposits and equity. In this sense, firms and banks are engaged in a security design problem and the result is a large array of potential securities that need to be priced and traded. We assume that competitive markets exist for all potential securities. In this sense, markets for *potential*

securities are *complete*.

Two non-standard assumptions are crucial for our results. The first is that default is costly. We model this very simply by assuming that, in the event of default, a bank or firm loses a fraction of its value. If financial distress were not costly, default would be a matter of indifference. The second assumption is that bank deposits yield a liquidity premium because they circulate as “money.” We model this very simply by assuming that consumers get extra utility if they pay for consumption with deposits. Without this assumption, there would be no need for banks. More precisely, the liquidity premium allows banks to borrow “cheaply,” that is, the return on deposits is lower than the return on equity. This cheap funding makes it possible for banks to offer loans to firms at attractive rates that compensate for the risk of financial distress. Without the liquidity premium, there would be no demand for bank loans and no need for banks. Costly default and a liquidity premium on deposits are sufficient to ensure the Modigliani-Miller theorem does not hold and that capital structures “matter” for banks or firms. Furthermore, even though, for the reasons explained above, there is some substitutability between equity in firms and in banks, we will see that firms and banks have their own comparative advantage when it comes to using equity as a buffer against default.

The first contribution of this paper is to provide a benchmark model where the constrained efficient asset composition and equity capital of banks and firms is decentralized by a market system, where each firm and bank independently maximizes its market value. The challenge is casting the model in a competitive general equilibrium framework. We assume away the usual frictions (imperfect competition, asymmetric information, etc.) but the model is rich in pecuniary externalities. Nonetheless, with a large enough set of potential security markets, decentralization of the constrained-efficient allocation is possible. The assumptions needed are not innocuous, however, as we discuss more fully in Section 3. We use constrained efficiency to characterize the equilibrium level of equity capital in firms and banks and show how it depends on the technology shocks that impact the economy.

The second contribution is to show that in the special case of co-monotonic technologies, it is socially optimal to allocate all equity capital to the corporate sector. A set of technologies is *co-monotonic* if the individual productivities are monotonically increasing in the state of nature. In this case, bank equity is dominated by firm equity and banks hold *zero* equity capital. This result highlights the comparative advantage of firm equity in buffering systemic shocks: firm equity provides a buffer against default for both banks and firms because it absorbs risk at the source, where the shock first hits the system.

This leads to our third contribution: why do banks need to hold equity capital? Our result for co-monotonic technologies suggests that positive bank capital is only efficient if firms are subject to idiosyncratic productivity shocks. We don’t have general results because the presence of non-convexities makes it very difficult to get a complete characterization of

equilibrium outside the co-monotonic case. We show first that a positive level of bank equity is optimal in a one factor model, that is simpler in some respects than the environment considered in the rest of the paper but also closer to the literature. The result is then extended to the more general set-up of the paper, for some examples of technology shocks for which equilibria that can be solved explicitly.

The optimality of bank equity arises when the probability that any given technology receives a negative productivity shock is relatively small while the probability that *some* technologies receive a shock is relatively large. In that case, a bank can reduce its risk by diversifying its portfolio of loans across firms using different technologies, but only if it holds equity to buffer against these shocks. The bank has a comparative advantage relative to firms in providing an equity buffer against these shocks for two reasons. First, it only needs to hold a small amount of equity because, in any state, only a small fraction of its diversified portfolio is affected. Second, because there is often some type of firm that is defaulting, the bank's equity buffer is more likely to be required than a firms' equity buffer. We then identify conditions under which both banks and firms will issue equity as well as conditions under which only banks issue equity.

The rest of the paper is organized as follows. In Section 2 we present the economy and the competitive equilibrium notion. In Section 3 we show the welfare properties of equilibria. Section 4 examines the properties of banks' capital structure in equilibrium. First, Section 4.1 shows that bank capital is positive in the single factor model with idiosyncratic risk and diversified portfolios. Section 4.2 considers the case of co-monotonic technologies while Section 4.3 studies an environment where productivity shocks have also an idiosyncratic component. Section 5 concludes. Proofs are collected in the appendix.

1.1 Related literature

Under conditions provided by Modigliani and Miller (1958), capital structure is indeterminate and has no effect on the value of the firm. A large literature has grown up investigating the role of various factors, such as taxes, bankruptcy costs, term structure, seniority and incentive problems, in the choice of corporate capital structure (Brennan and Schwartz, 1978; Barnea, Haugen and Senbet, 1981; Kim, 1982; Titman, 1984; Dammon and Green, 1987; Titman and Wessels, 1988; Leland and Toft, 1996; Bradley, Jarrell, and Kim, 2011; Hackbarth and Mauer, 2012). The Modigliani-Miller Theorem does not hold in our model either: the optimal capital structure is determined by a trade-off between the costs of financial distress and the funding advantage of debt, due to the liquidity premium on deposits.

Evidence that the costs of default can be substantial for both banks and non-financial firms is found in James (1991), Andrade and Kaplan (1998), and Korteweg (2010). More

recent work suggests that earlier estimates may have understated the true costs of default (Almeida and Philippon, 2007; Acharya, Bharath and Srinivasan, 2007). The default costs in our model are deadweight losses, as distinct from the fire sale “losses” in Gale and Gottardi (2015), which are merely transfers of value to buyers.

The empirical literature on the relationship between a bank’s capital structure and its market value is not large. Flannery and Rangan (2008) examined changes in banks’ capital structure in the previous decade. Mehran and Thakor (2011) found a positive relationship between bank value and bank capital in a cross section of banks. Gropp and Heider (2010) found that the determinants of bank capital structure were similar to those of non-financial firms, although the levels of equity are different.

The theoretical work on bank capital structure is also quite limited. Diamond and Dybvig (1983) and Diamond (1984) show that deposits constitute the optimal form of funding to provide liquidity insurance to depositors or delegated monitoring for investors. Gale (2004) extends the Diamond-Dybvig model to include bank capital that provides additional risk sharing between risk neutral equity holders and risk averse depositors. The importance of the liquidity services provided by deposits has been emphasized, more recently, by Stein (2012) and De Angelo and Stulz (2015).

Van den Heuvel (2008) studies a quantitative model in which bank capital structure is determined by the trade-off between the moral hazard costs associated with risk shifting behavior and the liquidity services of bank debt, assuming that deposits yield direct utility benefits. DeAngelo and Stulz (2015) also highlight the liquidity premium earned by bank deposits, contrasting it with the costs of intermediation.

Gornall and Strebulaev (2017), henceforth GS, provide a quantitative analysis of a partial equilibrium model in which the capital structures of banks and the firms that borrow from them are endogenously determined, weighing the tax advantage of debt against the costs of default. They find that the value maximizing leverage in the banking sector is much higher than in the corporate sector. They attribute this to the fact that banks are less risky than firms. First, banks hold senior claims (debt), so the first loss falls on corporate shareholders. Second, banks reduce their portfolio risk by diversifying across firms. In the environment considered, these two factors are sufficient to produce realistic levels of bank capital.

There are some important differences between our model and GS. In GS, the use of debt—by banks and firms—is motivated by the tax hedge on interest income. This implies that the interaction between firms and banks centers around the allocation of the tax benefits, while in our set-up it operates primarily on the default risk dimension. Furthermore, as a result of the distortionary tax, the equilibrium allocation is inefficient. The inefficiency could easily be removed by changing the tax code to treat so debt and equity symmetrically. But in that case, banks would have no funding advantage and firms would be funded entirely by equity.

Another feature of the GS model is the assumption that banks hold symmetric portfolios consisting of loans diversified across all firms. As we show in our model, this is not always the optimal choice of banks, because of the non-convexities associated with the possibility of costly default.

Allen, Carletti and Marquez (2015), henceforth ACM, motivated the present paper. Like us, they assume that banks and firms can only issue debt/deposits and equity. They also assume that markets for deposits and equity are segmented: some consumers can only hold deposits, whereas others can hold equity. Depositors have lower outside options than equity investors, so in equilibrium the depositors receive a lower return than the equity investors. Equity is therefore an “expensive” source of funding.

ACM characterize the capital structures of banks and firms, when they are chosen jointly to maximize total expected surplus, taking as given the expected returns demanded by the holders of equity and deposits, respectively. This optimal contracting approach guarantees the efficiency of these choices. In our framework, by contrast, efficiency is a property of a decentralized competitive equilibrium when markets are open for all possible types of debt and equity.

ACM also derive the result that bank equity has zero value, for the special case where firms’ returns are perfectly correlated and uniformly distributed. In that case, the bank is simply a pass-through for the shocks affecting the firms’ returns and the bank will default only if the firms it lends to default. As a result, putting all the equity in the firms minimizes the probability of default for both banks and firms. Our analysis generalizes this result and shows that only when one allows for a greater degree of generality in the structure of the technology shocks one can understand the possible benefits of bank equity.

There is a large theoretical literature analyzing the role of equity in preventing risk shifting or asset substitution. Our assumption of symmetric information rules out this kind of incentive problem.

Our competitive equilibrium model is related to the literature on the theory of the firm in incomplete markets, developed by Diamond (1967), Ekern and Wilson (1974), Radner (1974), Drèze (1974), and Grossman and Hart (1979). In the earlier literature, firms are fully owned by shareholders and the equilibrium value of a firm is determined by the marginal valuations of its owners. For example, a firm that produces a vector of future outputs \mathbf{y} has market value

$$v = \frac{\nabla u_i(\mathbf{x}_i) \cdot \mathbf{y}}{\|\nabla u_i(\mathbf{x}_i)\|}$$

where \mathbf{x}_i is the shareholder’s consumption bundle and $\nabla u_i(\mathbf{x}_i)$ is the vector of marginal utilities. Our assumption of complete markets for debt and equity implies the existence of equilibrium prices for all possible securities, even those that are not traded in equilibrium.

A similar approach is found in Makowski (1983) and Allen and Gale (1988, 1991). An alternative to the complete markets approach is to assume that only traded securities are priced, but that firms have rational conjectures about the value a security would have if a small amount of it were introduced. This approach was used by Hart (1979), for example, and appears to give the same results as the complete markets approach under sufficiently strong regularity conditions.

The existence of intermediaries and the costs of default make the pricing of assets more complicated in our model than in a stock market economy. Because a firm's debt is held by banks and default can occur at the firm level, the bank level, or both, the value of a firm's debt will depend on banks' willingness to pay for it, which in turn depends on the banks' capital structure and the consumers' willingness to pay for the debt and equity of banks. Bisin, Clementi and Gottardi (2017) also study the pricing of securities when intermediaries are present.

2 A general equilibrium model of capital structure

There are two dates, indexed $t = 0, 1$, and a finite number of states of nature, indexed $s = 1, \dots, S$. The true state is unknown at date 0 and revealed at date 1. The probability of state s at date 0 is denoted by $\pi_s > 0$, for $s = 1, \dots, S$, and is assumed to be common knowledge.

There are two goods, a capital good and a consumption good. There is a fixed endowment of the capital good at date 0, which is used as the sole input to produce the consumption good at date 1.

2.1 Producers

We assume there is a large mass of producers who use the capital good to produce the consumption good. Producers can choose from a finite number of technologies, indexed $j = 1, \dots, n$, for producing the consumption good. Using technology j , one unit of capital at date 0 produces $A_{js} > 0$ units of consumption at date 1 in state s . We assume that each producer can use at most one of the n technologies. Because production is subject to constant returns to scale, we can focus without loss of generality on the case where each firm uses one unit of capital. The amount of capital invested in each technology is then equal to the number of firms using that technology.

In this environment, a firm's capital structure is determined by the face value of the debt it issues. The face value of the debt is denoted by ℓ and belongs to a large but finite set $L \equiv \{0, \ell_1, \dots, \ell_{\max}\}$, where $0 \leq \ell_1 < \dots < \ell_{\max}$. Without loss of generality, we can assume

that $\ell_{\max} \leq \max \{A_{js}\}$ because there is no advantage in issuing debt with a face value that can never be paid.

Because productivity shocks are the only source of uncertainty, the technology choices made by firms determine the level of risk in the economy, while their capital structure choices determine how this risk is distributed between debt and equity.

Firms are identical ex ante but they may differ in their choice of technology and capital structure. A firm's choice of face value of debt ℓ and technology j is referred to as the firm's *type*. The set of firm types is denoted by $F \equiv L \times N$, with generic element (ℓ, j) , where $N = \{1, \dots, n\}$ denotes the set of available technologies.

A firm of type $f = (\ell, j) \in F$ issues debt and equity. The payoff vectors of these assets, denoted, respectively, by $\mathbf{a}_f^d \in \mathbf{R}_+^S$ and $\mathbf{a}_f^e \in \mathbf{R}_+^S$, are uniquely determined by the firm's type $f = (\ell, j)$ as follows:

$$a_{fs}^d = \begin{cases} \ell & \text{if } A_{js} \geq \ell \\ \lambda_f A_{js} & \text{if } A_{js} < \ell \end{cases} \quad (1)$$

and

$$a_{fs}^e = \begin{cases} A_{js} - \ell_j & \text{if } A_{js} \geq \ell \\ 0 & \text{if } A_{js} < \ell \end{cases} \quad (2)$$

for any state s . The parameter $0 \leq \lambda_f < 1$ is the firm's *recovery ratio* in the event of default. In other words, the default costs are $1 - \lambda_f$ per unit of output. For generality, we allow the recovery ratio to depend on the firm's type f , but this is not necessary and, in most applications, λ_f is independent of f .

Firms choose their technology and capital structure to maximize their profits, which is equivalent to maximizing the firm's market value. Since firms are subject to constant returns to scale, profits must be zero in equilibrium. In other words, the market value of the securities issued by a firm is just enough to finance the purchase of capital goods. Types of firms that cannot earn a zero profit will not operate in equilibrium.

Securities issued by firms are sold on competitive markets. In line with our completeness assumption, there is a price at which the securities issued by each type of firm are traded. Prices are denoted by the vector $\mathbf{q}_F = (\mathbf{q}_F^d, \mathbf{q}_F^e) \in \mathbf{R}_+^F \times \mathbf{R}_+^F$, where \mathbf{q}_F^d is the subvector of debt prices and \mathbf{q}_F^e is the subvector of equity prices. The *market value* of a firm of type f then is $q_f^d + q_f^e$. We normalize the price of capital goods to be equal to one. Hence, in equilibrium, we have $q_f^d + q_f^e \leq 1$ for any $f \in F$ —otherwise the demand for capital goods would be unbounded—and only the firm-types that achieve zero profits, $q_f^d + q_f^e = 1$, will operate in equilibrium.

2.2 Bankers

There is a large mass of bankers. Each banker can set up a single bank to lend to producers. For simplicity, we assume banks do not purchase the firms' equity¹. Bankers have no resources of their own and issue equity and deposits in order to finance their lending to producers.

A bank's capital structure is determined by the level of deposits it chooses to issue. The face value of deposits is denoted by d and is assumed to belong to a finite set $D \equiv \{0, d_1, \dots, d_{\max}\}$. The bank's portfolio is described by a vector $\mathbf{x} = \{x_f\}_{f \in F}$, where x_f is the *number* of loans made to firms of type f . Because the banks' technology is subject to constant returns to scale, we can, without loss of generality, focus on the case where each bank makes one unit of loans, that is.² $\sum_{f \in F} x_f = 1$. In other words, the bank invests in loans to firms that corresponds to the funding of one unit of capital goods³. As a consequence, a typical loan to a firm will be syndicated among several banks and each bank will hold a non-negative fraction of the loan. Let $X \subset R_+^F$ denote the set of admissible bank portfolios, also finite.

Each banker has access to the debt of all types of firms and has the same funding opportunities. Portfolios and capital structures may differ across banks, however. We refer to a bank's portfolio \mathbf{x} and capital structure d as its *type*. The set of bank types is $B = X \times D$. Let \mathbf{x}_b denote the portfolio of a bank of type b . The payoff vectors of the deposits and equity issued by a bank of type b are denoted by $\mathbf{a}_b^d \in \mathbf{R}_+^S$ and $\mathbf{a}_b^e \in \mathbf{R}_+^S$, respectively, and defined by

$$a_{bs}^d = \begin{cases} d & \text{if } \mathbf{x}_b \cdot \mathbf{a}_{Fs}^d \geq d \\ \lambda_b (\mathbf{x}_b \cdot \mathbf{a}_{Fs}^d) & \text{if } \mathbf{x}_b \cdot \mathbf{a}_{Fs}^d < d \end{cases} \quad (3)$$

and

$$a_{bs}^e = \begin{cases} \mathbf{x}_b \cdot \mathbf{a}_{Fs}^d - d & \text{if } \mathbf{x}_b \cdot \mathbf{a}_{Fs}^d \geq d \\ 0 & \text{if } \mathbf{x}_b \cdot \mathbf{a}_{Fs}^d < d, \end{cases} \quad (4)$$

for every state s , where the vector \mathbf{a}_{Fs}^d is defined by

$$\mathbf{a}_{Fs}^d = (\mathbf{a}_{fs}^d)_{f \in F}$$

¹There are differences between firms' debt and equity that might explain why banks do not invest in equity, or not as much as in debt: equity does not have a finite maturity, does not pay regular interest, and has a fluctuating value. Our two period model, based on a competitive equilibrium approach, where the set of existing markets is exogenously given, is not well suited for studying this and hence to provide microfoundations for our assumption. This assumption is standard in the banking literature, and makes it easier to compare our model to this literature.

²Given this property, the upper bound on the face value of banks' debt follows - also w.l.o.g. - from the one on firms' debt.

³The value of a loan will never be greater than the maximal face value of firm debt and we can set $d_{\max} = \ell_{\max}$ without loss of generality.

for each s . The recovery rate $0 \leq \lambda_b \leq 1$ is a constant and may or not depend on the bank's type $b \in B$.

The problem of each bank is to select its portfolio and its capital structure to maximize its profits, given by the difference between its market value, that is the value of the liabilities it issued, and the value of the portfolio it acquired. The bank takes as given the price of all debt claims issued by firms, $\mathbf{q}_F^d \in \mathbf{R}_+^F$, as well as the prices of all types of securities the banks can issue, $\mathbf{q}_B = (\mathbf{q}_B^d, \mathbf{q}_B^e) \in \mathbf{R}_+^B \times \mathbf{R}_+^B$, where \mathbf{q}_B^d is the subvector of deposit prices and \mathbf{q}_B^e is the subvector of equity prices. More formally, each bank will choose its type $b \in B$ to maximize market value minus the cost of the assets it acquired, $q_b^d + q_b^e - \mathbf{q}_F^d \cdot \mathbf{x}_b$. In equilibrium, the maximum profit will be zero, that is, $q_b^d + q_b^e \leq \mathbf{q}_F^d \cdot \mathbf{x}_b$, and only banks that achieve zero profits, $q_b^d + q_b^e = \mathbf{q}_F^d \cdot \mathbf{x}_b$, will be active.

2.3 Consumers

There is a unit mass of identical consumers. Each consumer has an initial endowment of one unit of capital at date 0 and no initial endowment of the consumption good. All consumption takes place at date 1, when the output of consumption good is realized. Consumers have VNM preferences over consumption described by

$$\sum_s \pi_s u(c_{1s} + \beta c_{2s}), \quad (5)$$

where c_{1s} denotes the consumption in state s that can occur immediately, because it is paid for with deposits, while c_{2s} denotes the consumption which is paid for with the yields of equity, and occurs with some delay. The constant $0 < \beta < 1$ captures the cost of this delay. The specification of the preferences reflects the assumption that deposits serve as money, whereas equity does not. The delay (or equivalently, transaction) costs involved in converting equity into “cash” are measured by the parameter $\beta \in (0, 1)$.⁴ The function $u : \mathbf{R}_+ \rightarrow \mathbf{R}$, describing the utility of total consumption in any state s , $c_{1s} + \beta c_{2s}$, is assumed to be increasing, concave and continuously differentiable.

Each consumer can use the revenue obtained by selling his endowment of capital at date 0 to purchase the deposits and equity issued by banks and the equity issued by firms. Consumers cannot purchase firm debt directly, but hold it indirectly by investing in banks that purchase firm debt.⁵ The pattern of trades in the economy is summarized in the following figure:

⁴The specification is a reduced-form representation of the greater convenience of using deposits for consumption compared to equity. A shareholder who wants to convert shares into consumption must pay a commission to sell the shares. Dividends are paid infrequently and must be converted into deposits before

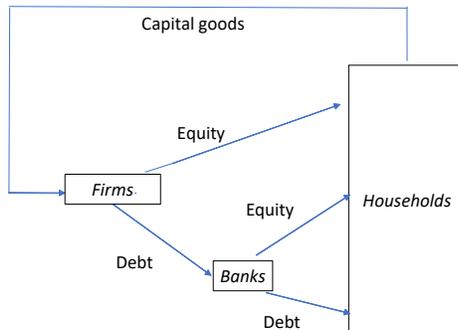


Figure 1: Trade flows in the economy

A consumer's portfolio is described by a vector $\mathbf{z} \equiv (\mathbf{z}_F, \mathbf{z}_B) \in \mathbf{R}_+^F \times \mathbf{R}_+^F \times \mathbf{R}_+^B \times \mathbf{R}_+^B$, where (z_b^d, z_b^e) denotes the consumer's demand for debt and equity issued by banks of type b and similarly for (z_f^d, z_f^e) . The set of feasible portfolios is denoted by Z and defined to be the set of portfolios \mathbf{z} such that $z_f^d = 0$ for all $f \in F$. Letting $\mathbf{q} \equiv (\mathbf{q}_F, \mathbf{q}_B)$, the consumer chooses a consumption bundle $\mathbf{c} = (\mathbf{c}_1, \mathbf{c}_2) \in \mathbf{R}_+^S \times \mathbf{R}_+^S$ and a portfolio $\mathbf{z} \in Z$ to maximize

$$U(\mathbf{c}) \equiv \sum_{s=1}^S \pi_s u(c_{1s} + \beta c_{2s})$$

they can be spent. This time delay reduces the value of the consumption because of discounting.

⁵This assumption allows us to simplify the analysis which follows, but could be relaxed without affecting the results. In the environment considered, even if consumers could hold firms' debt, deposits would be more attractive to them than firms' debt and banks would still have an advantage in lending to firms – because of their cheap financing via deposits. Hence in equilibrium banks would always hold at least some of the firms' debt. Under our assumption that the lowest shock hitting any of the firms' technologies is strictly positive, whatever the level of firms' debt banks can in fact offer safe debt, which earns a liquidity premium, without incurring any bankruptcy cost. Hence at the margin, starting from zero debt, banks would be willing to pay more than consumers for firms' debt.

Also, one may argue the assumption is quite realistic. In actual markets, households can hold corporate debt, but they typically do so by buying corporate bonds, not by lending directly to firms, usually a prerogative of banks who can monitor them. Bonds typically have large denominations and are traded, if at all, on illiquid markets. They do not serve as money.

subject to the budget constraints,

$$\begin{aligned} \mathbf{q} \cdot \mathbf{z} &\leq 1, \\ \mathbf{c}_1 &= \sum_{b \in B} z_b^d \mathbf{a}_b^d, \\ \mathbf{c}_2 &= \sum_{b \in B} z_b^e \mathbf{a}_b^e + \sum_{f \in F} z_f^e \mathbf{a}_f^e. \end{aligned}$$

2.4 Equilibrium

An allocation is described by a consumption bundle, \mathbf{c} , and a portfolio, \mathbf{z} , of the representative consumer, a distribution of banks over the set of possible bank types $\boldsymbol{\mu} = (\mu_b)_{b \in B}$, and a distribution of firm types $\boldsymbol{\kappa} = (\kappa_f)_{f \in F}$. Thus μ_b and κ_f are, respectively, the mass of bankers and producers who set up a bank of type b and a firm of type f . Formally, the *allocation* is an array $(\mathbf{c}, \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\kappa})$. An allocation is *attainable* if

$$\sum_{f \in F} \kappa_f = 1, \quad (6)$$

$$\sum_{b \in B} \mu_b \mathbf{x}_b = \boldsymbol{\kappa}, \quad (7)$$

$$z_b^d = z_b^e = \mu_b, \quad \forall b \in B, \quad (8)$$

$$z_f^e = \kappa_f, \quad \forall f \in F, \quad (9)$$

and

$$\mathbf{c} = \mathbf{z} \cdot \mathbf{a} = \left(\sum_{b \in B} z_b^d \mathbf{a}_b^d, \sum_{b \in B} z_b^e \mathbf{a}_b^e + \sum_{f \in F} z_f^e \mathbf{a}_f^e \right). \quad (10)$$

The first attainability condition (6) says that the firms collectively use the entire one unit of the capital good in the consumers' endowments⁶. The second condition, (7), says that banks hold in their portfolio all the debt issued by firms. The third and fourth conditions, (8) and (9), say that consumers hold all the deposits and equity issued by banks and all the equity issued by firms. Finally, the last condition, (10), restates the relationship between consumption and the payoff of the portfolio held by consumers.

An *equilibrium* consists of an attainable allocation $(\mathbf{c}, \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\kappa})$ and a price system \mathbf{q} such that:

- (i) $\kappa_f > 0$ only if f solves the firm's problem, given the prices \mathbf{q} ;

⁶Since each firm is assumed to use one unit of capital, the amount of capital used by firms of type f is also κ_f .

- (ii) $\mu_b > 0$ only if b solves the bank's problem, given the prices \mathbf{q} ;
- (iii) (\mathbf{c}, \mathbf{z}) solves the consumer's problem, given the prices \mathbf{q} .

Note that equilibrium condition (i) is equivalent to

$$\kappa_f > 0 \implies q_f^d + q_f^e = \max_{f \in F} \{q_f^d + q_f^e\} = 1,$$

for any $f \in F$. Similarly, equilibrium condition (ii) is equivalent to

$$\mu_b > 0 \implies q_b^d + q_b^e - \mathbf{q}_F^d \cdot \mathbf{x} = \max_{b \in B} \{q_b^d + q_b^e - \mathbf{q}_F^d \cdot \mathbf{x}\} = 0,$$

for any $b \in B$. In what follows, we refer to a firm of type f (respectively, bank of type b) as being *active* in equilibrium if and only if $\kappa_f^* > 0$ (respectively, $\mu_b^* > 0$). Also, prices are such that markets for the securities of non active firms clear with zero trades.

3 Efficiency of equilibria

In this section, we show that analogues of the fundamental theorems of welfare economics hold for the environment described above. Because banks and firms are restricted to using - non contingent - debt and equity, the appropriate welfare concept is not Pareto efficiency⁷, but *constrained* Pareto efficiency. We then show that all competitive equilibria⁸ are constrained efficient and any constrained efficient allocation can be decentralized as an equilibrium.⁹

We say that an attainable allocation $(\mathbf{c}^*, \mathbf{z}^*, \boldsymbol{\mu}^*, \boldsymbol{\kappa}^*)$ is *constrained Pareto efficient*, or constrained efficient, for short, if there does not exist an attainable allocation $(\mathbf{c}, \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\kappa})$ such that $U(\mathbf{c}) > U(\mathbf{c}^*)$. Formally, this is the case if and only if $(\mathbf{c}^*, \boldsymbol{\mu}^*)$ solves the problem

$$\max_{(\mathbf{c}, \boldsymbol{\mu})} \sum_{s=1}^S \pi_s u(c_{1s} + \beta c_{2s})$$

subject to the constraints

$$\sum_{b \in B} \mu_b = 1, \tag{11a}$$

⁷This restriction is typically binding, when firms' technologies are risky, hence the first best or Pareto efficient allocation cannot be attained. Firms and banks face then a trade-off between issuing liquid assets and costly default.

⁸In the online appendix we also establish, for completeness, the existence of a competitive equilibrium.

⁹Since there is a representative consumer, lump sum taxes and transfers are not needed to decentralize the constrained efficient allocation.

$$\mathbf{c} = \sum_{b \in B} \mu_b (\mathbf{a}_b^d, \mathbf{a}_b^e) + \sum_{b \in B} \mu_b \mathbf{x}_b \cdot \mathbf{a}_F^e. \quad (12a)$$

To see this, note first that if $(\mathbf{c}^*, \boldsymbol{\mu}^*)$ satisfies the constraints (11a) and (12a), we can use the attainability conditions (8) and (9) to define the consumers' portfolio \mathbf{z} and use the attainability condition (7) to define κ_f . Then it is easy to check that $(\mathbf{c}^*, \mathbf{z}^*, \boldsymbol{\mu}^*, \boldsymbol{\kappa}^*)$ satisfies the attainability constraints (6)–(10). Conversely, if $(\mathbf{c}^*, \mathbf{z}^*, \boldsymbol{\mu}^*, \boldsymbol{\kappa}^*)$ is an attainable allocation, $(\mathbf{c}^*, \boldsymbol{\mu}^*)$ satisfies the constraints.¹⁰

Proposition 1 *Let $(\mathbf{c}^*, \mathbf{z}^*, \boldsymbol{\mu}^*, \boldsymbol{\kappa}^*, \mathbf{q}^*)$ be an equilibrium. Then $(\mathbf{c}^*, \mathbf{z}^*, \boldsymbol{\mu}^*, \boldsymbol{\kappa}^*)$ is constrained Pareto efficient.*

Also, note that the set of attainable consumption vectors satisfying (11a), (12a) is convex and this allows us to establish the following:

Proposition 2 *Suppose that $(\mathbf{c}^*, \mathbf{z}^*, \boldsymbol{\mu}^*, \boldsymbol{\kappa}^*)$ is a constrained efficient allocation. Then there exists a price vector \mathbf{q}^* such that $(\mathbf{c}^*, \mathbf{z}^*, \boldsymbol{\mu}^*, \boldsymbol{\kappa}^*, \mathbf{q}^*)$ is an equilibrium.*

The argument of the proof of the two propositions is standard, and exploits the fact that in the competitive equilibrium notion, defined in Section 2.4, a market exists for all the possible types of securities that can be issued by firms and banks. The model considered provides a benchmark, where the risk composition of the assets and the capital structures chosen by firms and banks are efficient and we can then focus on the characterization of their properties. This benchmark is important, as it demonstrates the efficiency of firms' and banks' choices in a decentralized economy. Furthermore, it considerably simplifies the analysis of equilibrium allocations. For the validity of these results we assume away usual frictions, like asymmetric information, and some of the other conditions needed are also fairly demanding, as we discuss in what follows. As discussed in Section 5, our analysis provides useful objectives for regulatory interventions when such frictions are present.

First, implicit in the modeling of markets, is the assumption that agents know what securities they are trading. When a consumer makes a deposit in a bank, he knows the bank's capital structure and portfolio and that the bank is committed to these choices. Similarly, when a bank lends to a firm, it knows the firm's capital structure and portfolio and that the firm is committed to these choices. This assumption rules out risk shifting and other forms of moral hazard.

¹⁰To show this, we simply need to use the attainability conditions (7), (8), and (9) to eliminate \mathbf{z}^* and $\boldsymbol{\kappa}^*$ from (10), getting constraint (12a) as a result. Similarly, the attainability conditions (6) and (7) imply constraint (11a).

Second, as Geanakoplos and Polemarchakis (1986) have shown, competitive equilibrium is constrained inefficient for generic economies with uncertainty *if they satisfy certain conditions regarding the number of consumption goods and/or of trading dates*. In our framework, we assume there are only two dates and a single good at each date, so these conditions are not satisfied. The assumptions of a single good and two periods are common in financial applications, but they are restrictive. The representative consumer assumption is not crucial—as long as there is a single good, we could extend the theory to allow for multiple types of consumers.

Finally, as mentioned, we require the existence of markets (and market-clearing prices) for all *possible* types of debt and equity, not just the types traded in equilibrium. As we have already explained, firms and banks are creating new securities when they choose their capital structures and technologies (in the case of firms) or portfolios (in the case of banks). There is no limit on the number of possible securities, which could in principle be very large.

Since only a small subset of securities are actively traded in equilibrium, the prices in the inactive markets should be viewed as traders' expectations over the price at which newly issued securities would trade. It may be helpful to think about this assumption in terms of an analogy with the familiar idea of Arrow securities and their role in economizing on the number markets required to implement an efficient allocation. Instead of trading *all* contingent commodities at an initial date, agents trade Arrow securities at the initial date in order to reallocate their wealth across the uncertain states of nature. Then they trade commodities on spot markets, after the state of nature has been revealed. This two-step procedure reduces the number of markets in operation from $S \times L$ to $S + L$, where S is the number of states and L the number of physical commodities. But in order to economize on markets in this way, it is necessary for agents to anticipate the correct market-clearing prices in *all* the possible spot markets, those markets never open because the corresponding state does not occur, as well as the spots markets on which trade occurs. In the same way, the banks and firms in our model anticipate market-clearing prices for all securities markets, even those that are inactive.

4 Equilibrium capital structures

In this section we examine the interaction between the capital structure chosen by firms and the one chosen by banks. The use of equity is expensive because it reduces the portion of the cash flow that is paid out as liquid deposits. At the same time, banks and firms have incentives to include equity in their capital structures because equity provides a buffer against costly default. To some extent, firm equity and bank equity are substitutes, but

a firm's equity does "double duty" because it makes the firm's debt less risky and that in turn reduces the riskiness of the bank's assets. If a firm has very low leverage, its debt will be relatively safe. If a bank lends only to firms with low leverage, the bank's assets will be relatively safe and the need for an equity buffer will be diminished. This suggests that it may be efficient for banks to have higher leverage (lower equity buffers) than firms. At the same time, we will also see that bank equity provides a more effective buffer against idiosyncratic shocks.

The capital structure is not the only factor affecting the risk of default. The firms' choice of a technology also contributes to the level of aggregate risk in the economy and to the probability of default for individual firms and banks. Similarly, the banks' portfolio choices and the possibility of diversification may reduce the banks' probability of default.

Our analysis will show the comparative advantages of bank equity and firm equity. A general characterization of the equilibrium capital structures in the environment described is difficult because of the non-convexities introduced by the use of debt and the deadweight costs of default. We begin in the next subsection by considering a single factor model and showing conditions under which bank equity must be positive.

4.1 A "single factor" model

To explain the benefits of bank equity, based on its role in providing a buffer against *idiosyncratic* shocks, we present first the argument in the set-up of a "single factor" model. This has the advantage that the argument is simpler and the model is closer to the ones considered in the banking literature.

In the banking literature, single factor models, such as the one due to Vasicek (2002), are widely used to represent loss distributions of bank portfolios. This model considers a single bank that lends to an infinite number of firms indexed by j . Each borrowing firm j has the same marginal distribution of returns. The value of the assets $A_j(T)$ of firm j at time T are assumed to be the product of two random variables,

$$\log A_j(T) = \log A_j(0) + \log \sqrt{\rho} z + \log \sqrt{1 - \rho} \varepsilon_j, \quad (13)$$

where z is a standard normal aggregate shock, ε_j is a standard normal idiosyncratic shock to firm j , and ρ is the correlation parameter of the random variables $\{\varepsilon_j\}$. Firm j defaults at time T if $B_j > A_j(T)$, where B_j is the face value of the loan given by the bank.

The single factor model differs from the general equilibrium model presented in the previous section in several ways. The returns $A_j(T)$ are exogenous and have identical marginal distributions. The bank is perfectly diversified, that is, it lends the same amount to each firm j , and chooses its capital structure to maximize its value assuming risk neutrality.

Something similar to the specification in (13), when $\rho \in (0, 1)$,¹¹ can be obtained in the environment described in the previous section, if we assume that there is a large number of technologies $j \in [0, 1]$. The productivity of technology j (per unit of capital) is given by $A_j = \theta_j a$, where θ_j and a are random variables, with $\{\theta_j\}$ i.i.d. and independent of a . Let θ and a be distributed according to the c.d.f.s $F(\theta)$ and $G(a)$, respectively, where F and G are C^1 , both have support $[0, \infty)$, and the respective probability density functions $f(\theta)$ and $g(a)$ are strictly positive for $\theta > 0$ and $a > 0$. Furthermore, it is assumed that each technology is operated in the same proportion by firms and all firms have the same face value of the debt $\ell \geq 0$. Also, all banks hold a perfectly diversified and symmetric portfolio containing an equal amount of the debt issued by firms using each technology.

This single factor model differs from the general equilibrium model presented in the previous section because the main focus is on the capital structure decision of the representative bank, set to maximize its value when the representative consumer is risk neutral. Hence it abstracts from the decisions by firms and banks concerning the composition of their assets and from the possibility that firms may vary their debt level according to the technology chosen.¹² In this case, the return to a fully diversified portfolio depends only on the aggregate shock a . Let $\ell(a)$ be the amount repaid when the aggregate shock is a . Then

$$\ell(a) = \int_0^{\ell/a} \lambda_f \theta a f(\theta) d\theta + (1 - F(\ell/a)) \ell,$$

for any value of a . Under the assumptions about the distributions of a and θ , $\ell(a)$ is increasing in a . In fact,

$$\begin{aligned} \ell'(a) &= (\lambda_f \ell f(\ell/a) - \ell f(\ell/a)) \frac{-\ell}{a^2} + \int_0^{\ell/a} \lambda_f \theta f(\theta) d\theta \\ &= (1 - \lambda_f) \left(\frac{\ell}{a}\right)^2 f(\ell/a) + \int_0^{\ell/a} \lambda_f \theta f(\theta) d\theta > 0. \end{aligned}$$

The return to bank deposits, with face value \bar{d} , when the aggregate shock is a , is denoted by $d(a)$ and defined by

$$d(a) = \begin{cases} \bar{d} & \text{if } \bar{d} \leq \ell(a) \\ \lambda_b \ell(a) & \text{if } \bar{d} > \ell(a). \end{cases}$$

¹¹The two extreme cases, where returns are purely idiosyncratic, $\rho = 0$, or perfectly correlated, $\rho = 1$, are ignored here as they are not very interesting situations for the choice of the banks' capital structure. In both situations, as explained in the next sections 4.2 and 4.3, banks do not need to hold any capital.

¹²The properties we postulated for these decisions appear natural in the environment we described, because the technologies are ex ante identical and subject to some shocks that are perfectly diversifiable. In a later section, we demonstrate, for a simpler but analogous version of the technology shocks, that all these properties arise endogenously in equilibrium.

The return to bank equity is $\max\{\ell(a) - \bar{d}, 0\}$, while the return to firm equity, if the aggregate shock is a , is

$$\int_{\ell/a}^{\infty} (\theta a - \ell) f(\theta) d\theta.$$

An efficient choice of ℓ and \bar{d} must maximize the representative agent's utility, assumed risk neutral. This can be expressed in terms of the payoffs of deposits and bank and firm equity specified above:

$$\int_0^{\infty} \left\{ d(a) + \beta \max\{\ell(a) - \bar{d}, 0\} + \beta \int_{\ell/a}^{\infty} (\theta a - \ell) f(\theta) d\theta \right\} g(a) da. \quad (14)$$

Since the yield of firms' debt is less or equal than its face value ℓ and default is costly, we must have $\bar{d} \leq \ell$ in equilibrium. The expected returns to bank equity will be positive if and only if $\bar{d} < \ell(a)$ for a set of values of a with positive probability. But this must always be true if $\bar{d} < \ell$. To see this, note that if $\ell > 0$, we have $\ell(a) < \ell$ for any finite value of a and $\ell(a) \rightarrow \ell$ as $a \rightarrow \infty$. Since $\ell(a)$ is continuous, if $\bar{d} < \ell$ there is a positive probability that $\bar{d} < \ell(a)$ and the value of the bank's equity is positive.

Proposition 3 *In the single factor model, under the maintained assumptions, banks choose a positive value of equity: the values $\hat{\bar{d}}$ and $\hat{\ell}$ that maximize (14) must satisfy $\hat{\bar{d}} < \hat{\ell}$.*

Why does the efficient bank capital structure have a positive value of equity in the single factor model? In this environment, increasing the value of bank equity above zero reduces the bankruptcy costs of banks, by an amount $(1 - \lambda_b) \ell(a)$, by avoiding default for high realizations of a . The substitution of deposits with equity also entails a cost, equal to $(1 - \beta) (\ell(a) - \bar{d})$, due to the decrease in the payout to depositors. We show in the proof that, provided bank equity is not too large, the first effect always prevails over the second.

The result crucially depends on the existence of the idiosyncratic shock θ , which ensures that the return $\ell(a)$ on the bank's portfolio of loans is always strictly increasing in a . If θ were a degenerate random variable, for example, $\theta = 1$ with probability one, this property is no longer valid. In fact, $\ell(a)$ would be constant for sufficiently large values of a , and the constrained efficient value of equity would be zero. This result is shown formally in the next section, for the general equilibrium model described in Section 2.

4.2 Co-monotonic technologies

In this section, we analyze the properties of the equilibria, where each firm chooses a technology and its leverage, and each bank its loan portfolios and leverage, under the assumption that the n available technologies are *co-monotonic*.

Definition 4 *Technologies are said to be co-monotonic if $A_{js-1} < A_{js}$, for every $s = 2, \dots, S$ and $j = 1, \dots, n$.*

This condition requires that the productivities of all technologies are increasing functions of the state s . In other words, the productivity shocks are driven by a single factor and there is no idiosyncratic component. As a consequence, an increase in s reduces defaults for all types of firms and banks. If the co-monotonicity property is satisfied and the sets L and D are sufficiently rich in a sense we make precise in the proof, we obtain the following result, which generalizes the one by ACM:¹³

Proposition 5 *Assume that technologies are co-monotonic. Then if $(c^*, z^*, \mu^*, \kappa^*, q^*)$ is an equilibrium, the value of bank equity is zero for all active bank types $b \in B$.*

Thus, in equilibrium, banks default if any of their borrowers default. Each bank is on a knife edge, with no capacity to absorb losses. The result shows that the “double duty” role of firm equity prevails here. In any configuration where banks have positive equity, there must be a state where banks are solvent but hold in their portfolio the debt of a defaulting firm. In that case, an improvement can be achieved by decreasing the face value of the debt of such firm and increasing its equity, so that the firm becomes solvent in this state. Doing so reduces the firm’s default costs and hence increases the total payment to equityholders, without affecting payouts to depositors.

The co-monotonicity assumption is stated as a property of the productivity of all the technologies available in the economy. It is easy to see from the proof of Proposition 5 that this result is valid as long as the bank lends only to firms with co-monotonic technologies. For any bank portfolio \mathbf{x} , let the set of technologies represented in the portfolio be denoted by $J(\mathbf{x})$ and defined by

$$J(\mathbf{x}) = \{j = 1, \dots, n : x_{(\ell,j)} > 0 \text{ for some } \ell\}.$$

Then we say that the portfolio \mathbf{x} is *co-monotonic* if the set of technologies $J(\mathbf{x})$ is co-monotonic in the usual sense. The following corollary is then immediate.

Corollary 6 *In any equilibrium $(c^*, z^*, \mu^*, \kappa^*, q^*)$, the value of equity is zero for any active bank b^* whose portfolio \mathbf{x}^* is co-monotonic.*

¹³ACM make a number of other restrictive assumptions not required in our framework: consumers are risk neutral and exogenously divided into depositors and shareholders; there is a single technology with uniformly distributed productivity; the capital structure of banks and firms is chosen to maximize total surplus.

ACM establish an analogous result when a single technology is available to firms, a special case of co-monotonicity, under some assumptions recalled in the Introduction. They also extend the result to an environment with two technologies with i.i.d. productivities. The reason is that under appropriate conditions, because of the non-convexities generated by the possibility of costly default, it is optimal for banks not to diversify their portfolio and to lend only to firms that use the same technology (in other words, banks choose co-monotonic portfolios as in Corollary 6). We will show in the next section that this last property is not robust and, with enough independent technologies, banks choose to diversify their portfolio; when not all risk can be diversified away – as in the single factor model – we have then positive bank equity.

4.3 Positive bank equity

The results in the previous section show that the presence of an idiosyncratic component in the technology shocks is necessary for bank equity to have positive value, but is it sufficient outside the assumptions of the single factor model? To answer this question, in this section we study a simple specification of the technology shocks, with both aggregate and idiosyncratic productivity shocks, for which we can derive the equilibrium allocation and capital structures explicitly. We can then identify conditions under which competitive equilibria exhibit a positive value of bank equity.

We assume there are n technologies and a finite number $S = n + 2$ of states of nature. The probability of state $s \in \{1, \dots, n + 2\}$ is denoted by π_s and given by

$$\pi_s = \begin{cases} \frac{1-\delta-\varepsilon}{n} & \text{for } 1 \leq s \leq n \\ \delta & \text{for } s = n + 1 \\ \varepsilon & \text{for } s = n + 2 \end{cases} . \quad (15)$$

The productivity of technology $j \in \{1, \dots, n\}$ in state s is assumed to satisfy

$$A_{js} = \begin{cases} a_L & \text{if } s = j \leq n \\ a_M & \text{if } s \neq j \leq n \\ a_M & \text{if } s = n + 1 \\ a_H & \text{if } s = n + 2 \end{cases} , \quad (16)$$

where $0 < a_L < a_M < a_H$.

This specification features both aggregate and idiosyncratic shocks. It has been chosen so as to capture the different effects that these two components of the uncertainty have on firms' and banks' choices, while keeping the characterization of equilibria still tractable. The idiosyncratic shocks affect the technologies in states $s = 1, \dots, n$, in the sense that one and a

different technology in each of these states features a drop in its productivity to a_L , but the average productivity across technologies is constant in each of these states. In contrast, in states $n + 1$ and $n + 2$ respectively a small and big positive shock affect the productivity of all the technologies. (In the Appendix we discuss more in detail how this specification relates to the single factor model of Section 4.1.) Varying the parameters ε, δ , the relative weight of aggregate and idiosyncratic uncertainty varies, in different ways, and we can then analyze the effect on firm and bank choices.

We will also assume here that consumers are risk neutral, which allows us to simplify the characterization of equilibrium prices and quantities. Under this assumption, an attainable allocation $(\mathbf{c}^*, \mathbf{z}^*, \boldsymbol{\mu}^*, \boldsymbol{\kappa}^*)$ is constrained efficient if and only if any type of bank b^* in the support of $\boldsymbol{\mu}^*$ satisfies

$$\sum_{s=1}^S \pi_s (c_{1s}^* + \beta c_{2s}^*) = \sum_{s=1}^S \pi_s (a_{b^*s}^d + \beta (a_{b^*s}^e + \mathbf{x}_{b^*}^* \cdot \mathbf{a}_F^e)).$$

In other words, if we think of an active bank and the firms that borrow from it as a conglomerate, the market value of this conglomerate must equal the expected value of consumption for the representative consumer. If this condition were not satisfied, either the bank or the firms or both would not be maximizing their market values. We use this property repeatedly in what follows.

A bank portfolio is said to be *simple* if all firms whose debt is held by the bank have the same capital structure, as long as they have the same technology, or the same probability distribution of their productivity. We can show that it is optimal for banks to choose a simple portfolio when the structure of technologies satisfies (15) and (16), thus providing some justification for the conditions used in the analysis of the single factor model. Since the technologies are ex ante identical, it follows that the same capital structure is optimal for all firms.

Proposition 7 *Suppose the technologies satisfy (15) and (16). Then, in equilibrium, every bank chooses a simple portfolio, that is, one containing either (i) only the debt of firms with no default risk ($\ell_j = a_L$), or (ii) only the debt of firms with low default risk ($\ell_j = a_M$), or (iii) only the debt of firms with high default risk ($\ell_j = a_H$).*

Given the technology structure, there are three possible candidates for the face value of the firm's debt. The firm can choose the face value equal to a_L , so that it never defaults, or a_M , so that the firm only defaults when hit by a negative shock, or a_H , so that the firm defaults unless it is hit by a large positive shock. Proposition 7 assures us that a bank will

lend to firms that use only one of these capital structures.¹⁴

The specification given by (15) and (16) incorporates a number of interesting cases. In the limit as $\delta + \varepsilon \rightarrow 1$, that is, when only states $n + 1$ and $n + 2$ have positive probability, there is only aggregate risk and all technologies are identical and, hence, co-monotonic. Then Proposition 5 implies that the value of bank equity will be zero in any equilibrium when $\delta + \varepsilon = 1$. At the other extreme, in the limit as $\delta \rightarrow 0$ and $\varepsilon \rightarrow 0$, only the states $s = 1, \dots, n$ can occur and in each of them, exactly one technology yields a_L and the remainder yield a_M . We have so the case of pure idiosyncratic risk. In this case too, we show that bank equity has zero value, but for rather different reasons than in the co-monotonic case. The following proposition characterizes the equilibrium allocations in such a situation. These allocations are, of course, constrained efficient by Proposition 1.

Proposition 8 *Assume the technologies satisfy (15), (16) and $\delta = \varepsilon = 0$. Then in equilibrium banks' choices are as follows:*

- (i) *if $\frac{1}{1-\beta} \left(\frac{n-\lambda_f}{n-1} - \beta \right) a_L > a_M$, each bank lends exclusively to firms with safe debt ($\ell_j = a_L$), and sets its level of deposits at $d = a_L$,¹⁵ or*
- (ii) *if $\frac{1}{1-\beta} \left(\frac{n-\lambda_f}{n-1} - \beta \right) a_L < a_M$, each bank lends exclusively to firms with risky debt ($\ell_j = a_M$) and chooses a fully diversified portfolio (holding a fraction $1/n$ of debt of firms choosing technology j , for each $j \in \{1, \dots, n\}$) and sets*

$$d = \frac{n-1}{n} a_M + \frac{1}{n} \lambda_f a_L.$$

In either case bank equity has zero value.

The proposition shows that when $\delta = \varepsilon = 0$, there are two possible equilibrium allocations, distinguished by the face value of firms' debt and by whether or not firms default. In each equilibrium allocation, banks' portfolios are risk free. Note that a risk free portfolio should always be entirely funded with deposits because equity is costly ($\beta < 1$) and there is no need for an equity buffer in absence of default risk. Moreover, firms choose higher leverage, so that they default if hit by a negative shock, when the size of the negative idiosyncratic shock, $a_M - a_L$, is sufficiently large, and/or the liquidity $(1 - \beta)$ or the default costs $(1 - \lambda_f)$ are sufficiently small.

In both of the extreme cases in the environment considered in this section, that is, when $\delta = \varepsilon = 0$ and $\delta + \varepsilon = 1$, it is optimal to have zero bank equity. In between the two extremes, when there is both aggregate risk and idiosyncratic risk, there is a role for equity

¹⁴Unfortunately, there is no intuitive explanation for this result. The proof proceeds by considering all possible portfolios and showing that non-simple portfolios are always dominated.

¹⁵The choice technology is immaterial since safe debt pays a_L with probability one.

in the efficient capital structures of both firms and banks. The following result establishes this more precisely.

Proposition 9 *Assume the technologies satisfy (15) and (16). Then if*

$$\frac{1}{1-\beta} \left(\frac{n-\lambda_f}{n-1} - \beta \right) a_L < a_M,$$

there exist $\delta^ > 0$ and $\varepsilon^* > 0$ such that, for any $0 < \delta < \delta^*$ and $0 \leq \varepsilon < \varepsilon^*$, in any equilibrium, each bank chooses a fully diversified portfolio (consisting of equal amounts of loans to firms choosing each technology $j = 1, \dots, n$), with face value $\ell = a_M$. The face value of deposits is*

$$d = \frac{n-1}{n} a_M + \frac{1}{n} \lambda_f a_L,$$

and bank equity has positive value.

In the situation described, banks receive a loan repayment equal to $a_M > d$ in state $s = n+1, n+2$. Hence the value of bank equity is positive. The same is true also for firm equity if $\varepsilon > 0$, because the shareholders of the representative firm receive $a_H - a_M > 0$ in state $s = n+2$.

The first condition required for the claim in Proposition 9 is the one derived in case (ii) of Proposition 8 under which, when $\delta = \varepsilon = 0$, firms choose a high debt level $\ell = a_M$ and banks choose a fully diversified portfolio in equilibrium. The second condition is that δ and ε are both positive, but not too large. That is, we are sufficiently close to the case of purely idiosyncratic uncertainty so that firms' and banks' debt level and banks' portfolio are the same as when $\delta = \varepsilon = 0$.

To gain some understanding of why banks choose a positive level of equity under the conditions of Proposition 9, note that firms choose a risky face value of debt, $\ell_j = a_M$, so a firm choosing technology j defaults in state $s = j$. To avoid default in all states, the firm would have to reduce its debt level ℓ_j from a_M to a_L . Given the relatively small likelihood of state $s = j$, the cost would outweigh the benefits. The opposite is true if we assess costs and benefits of an equity buffer to prevent the default of a diversified bank. First, the probability that *some* firm whose debt is held by the bank defaults is n times higher than the probability that any given firm defaults. A shock that is unlikely to affect a particular firm is so quite likely to affect a diversified bank. Second, the required equity buffer is much smaller than in the firm's case, because the debt of the defaulting firm is a small fraction of the bank's portfolio. We begin so to see why banks choose to hold a capital buffer against a shock while firms do not.

The first three lines in the table below summarize the equilibrium values of bank and firm equity in the different cases characterized by Propositions 8 and 9:¹⁶

Parameters	Value of bank equity	Value of firm equity
$\varepsilon + \delta = 1$	$= 0$??
$\varepsilon + \delta = 0$	$= 0$	$= 0$
$0 \leq \delta < \delta^*, 0 < \varepsilon < \varepsilon^*$	> 0	> 0
$\delta^{**} < \delta < 1, \varepsilon = 0$	$= 0$	$= 0$

Note that whether firm equity has positive or zero value in the case where the only aggregate risk and technologies are co-monotonic ($\varepsilon + \delta = 1$) depends on parameter values. There are two possible equilibrium outcomes: in one $\ell_j = a_M$ and the value of firm equity is positive, in the other, $\ell_j = a_H$ and the value of firm equity is zero.

The technologies defined by (15) and (16) also allow us to explore the connection between portfolio diversification and the presence of positive bank equity. We noted at the end of Section 4.2 that bank equity cannot have positive value unless banks diversify their portfolios. In other words, bank equity has no role as a buffer unless the bank faces risks that can be partly diversified. At the same time, one can have diversification without a positive value of bank equity, as we saw in case (ii) of Proposition 8, where bank equity has zero value and the fully diversified bank's portfolio is risk free. But there are other situations (described in the last line of the table above) where there is a risk of default that could be reduced by diversification, yet banks choose not to diversify and to issue no equity, accepting a positive probability of default.

Proposition 10 *Assume that firms' technologies satisfy (15), (16) and $\varepsilon = 0$. Then there exists $\delta^{**} > 0$ such that, for all $\delta^{**} < \delta < 1$, any equilibrium has the following properties: each bank lends to firms using a single technology j and issuing debt with low default risk ($\ell_j = a_M$), and bank deposits are $d = a_M$. Both these firms and the banks lending to them default in state $s = j$.*

The proof follows an argument parallel to the proof of Proposition 9. In the limit, when $\delta = 1$, there is no risk and banks and firms choose $d = \ell = a_M$. For $\delta < 1$ sufficiently close to 1, the upper hemicontinuity of the constrained efficient allocation correspondence implies that Proposition 10 holds. To avoid banks' default, even with diversification, an equity buffer is needed, as we saw in Proposition 9, that is bounded away from zero. As $\delta \rightarrow 1$, the cost of the buffer remains bounded away from zero while the benefit converges to zero. When the bank chooses $d = a_M$, on the other hand, the value of diversification is negative: the

¹⁶Note that the condition $\frac{1}{1-\beta} \left(\frac{n-\lambda_f}{n-1} - \beta \right) a_L < a_M$ is maintained here.

probability of default is minimized by choosing a portfolio made of loans to firms using the same technology j and the benefit of this choice increases as $\delta \rightarrow 1$.

Thus, even if diversification is possible, it is not always optimal. Evidently, the gains from diversification have to be “sufficiently large” before they provide a motive for banks to issue a positive value of equity. In other words, a violation of co-monotonicity is necessary, but not sufficient, as we saw in Proposition 10, for bank equity to have positive value. This principle is also illustrated by ACM, who analyse an environment with two i.i.d. technologies and show that, for some parameter values, it is optimal for banks to lend only to firms using one of the technologies, that is, to choose a co-monotonic portfolio. In that case, it will be optimal for bank equity to have zero value, even though there appears to be scope for diversification.

5 Conclusion

We have presented a competitive equilibrium model in which banks act as intermediaries between consumers and productive firms that choose which risky technology to operate. Banks and firms raise funds by issuing debt and equity. The Modigliani Miller Theorem does not hold in this environment for two reasons. First, bank deposits offer liquidity services. Second, the use of debt can lead to costly bankruptcy. As a result, the optimal capital structure, for a firm or a bank, is determinate in equilibrium. Firms’ equity serves as a buffer against the default risk both of firms and of banks lending to them, while banks’ equity, together with the diversification of their portfolio, is only a buffer against banks’ default risk. Hence there is some trade-off between allocating capital buffers to firms rather than banks and to determine the level of bank equity the entire supply chain of financing must be taken into account.

In this context, we have established analogues of the fundamental theorems of welfare economics, showing that equilibria are constrained efficient and that constrained efficient allocations are decentralizable. Thus, equilibrium capital structures are privately and socially optimal: they maximize the market value of the firm or bank and they are consistent with constrained efficiency. The importance of the general equilibrium theory is that it shows how markets coordinate the choice of capital structures, in the corporate and banking sectors, together with the firms’ choice of technology risk, so that risk and equity capital are optimally allocated between the two sectors.

We have shown that the equilibrium value of bank equity depends on the risk structure of the economy. In particular, bank equity is zero when there is only aggregate risk. In contrast, bank equity is positive when there is both aggregate and idiosyncratic risk and sufficient scope

for diversification so that most – though not all – of the risk can be eliminated in banks.

As we mentioned in the Introduction, the study of an environment in which the laissez-faire equilibrium is constrained efficient provides a benchmark, and a useful starting point for a theory of bank capital regulation. In the first place, the characterization of the constrained efficient capital structures gives us some idea of which kind of allocations capital adequacy regulation should aim to achieve when equilibria fail to be efficient. In the second place, the contrast between the ideal conditions under which laissez-faire equilibria are constrained efficient and the ones present in the real world may allow us to gain some understanding of why market failures occur and also to assess their welfare costs.

We also alluded briefly to some of the frictions which may justify capital regulation. Asymmetric information is one. Banks are opaque and depositors and equity holders alike may have incomplete information about the risk characteristics of the bank's portfolio. Asymmetric information may give rise to moral hazard in the form of risk shifting and asset substitution. Furthermore, the possibility of government bailouts in the event of default may encourage excessive risk taking. Externalities, whether pecuniary or real, may give rise to costs that are not internalized by bankers.

Capital requirements may not be a panacea, however. Managers' interests are not aligned with shareholders' interests and, even if they are aligned, it is not clear that the top management of the largest banks is aware of and able to control risk taking by highly incentivized managers at lower levels. The financial crisis provided us with several examples of high level managers, holding large equity stakes in the bank, who were unaware of the dangers facing their banks until the last minute.

We have also ignored a second motive for requiring a large capital buffer: bank capital is part of total loss absorption capacity (TLAC), which reduces the need for politically unpopular bailouts. Whether innovations such as TLAC and bail-inable debt will actually put an end to bailouts is not clear, but capital regulation motivated by the desire to avoid bailouts may look quite different from capital regulation motivated by moral hazard in risk taking.

So, there is much to be done in order to develop a satisfactory microfoundation for capital regulation. But the recognition that efficient capital structures in the banking and corporate sectors are interrelated and determined by general equilibrium forces is a first-order requirement for any sensible theory.

References

- [1] Acharya, V., S. Bharath, and A. Srinivasan (2007): Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries. *Journal of Financial Economics*, **85**, pp. 787–821.
- [2] Allen, F., E. Carletti, and R. Marquez (2015): Deposits and Bank Capital Structure. *Journal of Financial Economics*, **118**, pp. 601-19
- [3] Allen, F. and D. Gale (1988): Optimal Security Design. *Review of Financial Studies*, **1**, pp. 229-263
- [4] Allen, F. and D. Gale (1991): Arbitrage, Short Sales, and Financial Innovation. *Econometrica*, **59**, pp. 1041-1068
- [5] Almeida, H. and T. Philippon (2007): The risk-adjusted cost of financial distress. *Journal of Finance*, **62**, pp. 2557–2586
- [6] Andrade, G. and S. Kaplan (1998): How costly is financial (not economic) distress? Evidence from highly levered transactions that became distressed. *Journal of Finance*, **53**, pp. 1443–1493
- [7] Barnea, A., R. A. Haugen, and L. W. Senbet (1981): An equilibrium analysis of debt financing under costly tax arbitrage and agency problems. *Journal of Finance* **36**:569–81.
- [8] Bisin, A., G. Clementi and P. Gottardi (2017): Capital Structure and Hedging Demand with Incomplete Markets, NYU Working Paper.
- [9] Bradley, M., G. A. Jarrell, and E. H. Kim (2011): On the existence of an optimal capital structure: Theory and evidence. *Journal of Finance*, **39**, pp. 857–78.
- [10] Brennan, M. J., and E. S. Schwartz (1978): Corporate income taxes, valuation, and the problem of optimal capital structure. *Journal of Business*, **51**, pp. 103–14.
- [11] Dammon, R. M., and R. C. Green (1987): Tax arbitrage and the existence of equilibrium prices for financial assets. *Journal of Finance*, **42**, pp. 1143–66.
- [12] DeAngelo, H., and R. Stulz (2015): Liquid-claim production, risk management, and bank capital structure: why high leverage is optimal for banks. *Journal of Financial Economics*, **116**, pp. 219-236

- [13] Diamond, P.A: (1967): The Role of a Stock Market in a General Equilibrium Model with Technological Uncertainty, *American Economic Review*, **57**, pp. 759-776.
- [14] Diamond, D. (1984): Financial intermediation and delegated monitoring. *Review of Economic Studies*, **60**, pp. 393–414.
- [15] Diamond, D., and P. Dybvig (1983): Bank runs, deposit insurance and liquidity. *Journal of Political Economy*, **91**, pp. 401–419.
- [16] Dreze, J.H: (1974): Investment under private ownership: Optimality, equilibrium and stability, in J.H. Dreze (ed.), *Allocation under Uncertainty: Equilibrium and Optimality*, New York, McMillan.
- [17] Ekern, S. and R. Wilson (1974): On the Theory of the Firm in an Economy with Incomplete Markets. *Bell Journal of Economics and Management Science*, **5**, pp. 171-180.
- [18] Flannery, M. and K. Rangan (2008): Market forces at work in the banking industry: evidence from the capital buildup from the 1990s. *Review of Finance*, **12**, pp. 391–429
- [19] Gale, D. (2004): Notes on optimal capital regulation. In: St-Amant, P., C. Wilkins (Eds.), *The Evolving Financial System and Public Policy*. Bank of Canada, Ottawa, Ontario, Canada, pp. 225–253.
- [20] Gale, D. and P. Gottardi (2015): Capital Structure, Investment and Fire Sales. *Review of Financial Studies*, **28**, pp. 2502-33.
- [21] Geanakoplos, J.D. and H.M. Polemarchakis (1986): Existence, Regularity, and Constrained Suboptimality of Competitive Allocations when the Asset Market is Incomplete, in Heller, W.P., R.M. Starr and D. Starrett (Eds), *Uncertainty, Information, and Communication: Essays in Honor of Kenneth J. Arrow*, Vol. III, Cambridge: Cambridge University Press, 65-95.
- [22] Gornall, W. and I.A. Strebulaev (2017): Financing as a Supply Chain: The Capital Structure of Banks and Borrowers, Stanford University Working Paper.
- [23] Gropp, R. and F. Heider (2010): The determinants of capital structure: some evidence from banks. *Review of Finance*, **14**, pp. 587–622.
- [24] Grossman, S.J. and O.D. Hart (1979): A theory of competitive equilibrium in stock market economies, *Econometrica*, **47**, pp. 293-329.

- [25] Hackbarth, D., and D. C. Mauer (2012): Optimal priority structure, capital structure, and investment. *Review of Financial Studies*, **25**, pp. 747–96.
- [26] Hart, O.D. (1979): On shareholder unanimity in stock market economies, *Econometrica*, **47**, pp. 1057-83.
- [27] James, C. (1991): The losses realized in bank failures. *Journal of Finance*, **46**, pp. 1223–1242
- [28] Kim, E. H. (1982): Miller’s equilibrium, shareholder leverage clienteles, and optimal capital structure. *Journal of Finance*, **37**, pp. 301–19.
- [29] Korteweg, A. (2010): The net benefits to leverage. *Journal of Finance*, **65**, pp. 2137–2170
- [30] Leland, H. E., and K. B. Toft (1996): Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *Journal of Finance*, **51**, pp. 987–1019.
- [31] Makowski, L. (1983): Competitive stock markets. *Review of Economic Studies*, **50**, pp. 305-330.
- [32] Mehran, H. and A. Thakor (2011): Bank capital and value in the cross section. *Review of Financial Studies*, **24**, pp. 1019–1067.
- [33] Modigliani, F. and M. Miller (1958): The cost of capital, corporation finance, and the theory of investment *American Economic Review*, **48**, pp. 261–297.
- [34] Radner, R. (1974): A note un unanimity of stockholders preferences among alternative production plans. A reformulation of the Ekern-Wilson model, *Bell Journal of Economics and Management Science*, **5**, pp. 181-184.
- [35] Stein, J.C. (2012): Monetary policy as financial stability regulation, *Quarterly Journal of Economics*, *127*, pp. 57-95.
- [36] Titman, S. (1984): The effect of capital structure on a firm’s liquidation decision. *Journal of Financial Economics*, **13**, pp. 137–51.
- [37] Titman, S., and R. Wessels (1988): The determinants of capital structure choice. *Journal of Finance*, **43**, pp.1–19.
- [38] Van den Heuvel, S. (2008): The welfare cost of bank capital requirements. *Journal of Monetary Economics*, **55**, pp. 298–320

[39] Vasicek, O. (2002): Loan Portfolio Value. *Risk*, **15**, pp. 160–62.

Appendix

Proof of Proposition 1

The proof is by contradiction. Suppose, contrary to what we want to prove, that $(\mathbf{c}^*, \mathbf{z}^*, \boldsymbol{\mu}^*, \boldsymbol{\kappa}^*)$ is constrained inefficient. Then there exists a feasible allocation $(\mathbf{c}, \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\kappa})$ such that $U(\mathbf{c}) > U(\mathbf{c}^*)$. Condition (iii) of the definition of equilibrium implies that \mathbf{c} lies outside the representative consumer's budget set. That is, $\mathbf{q}^* \cdot \mathbf{z} > 1$, where

$$\mathbf{c} = \mathbf{z} \cdot \mathbf{a} = \sum_{h \in H} (z_h^d \mathbf{a}_h^d, z_h^e \mathbf{a}_h^e)$$

and $H = B \cup F$ is the combined set of types of banks and firms. Now the equilibrium optimality condition for banks implies that

$$q_b^{d*} + q_b^{e*} \leq \mathbf{q}_F^{d*} \cdot \mathbf{x}_b,$$

for any $b \in B$. Then

$$\begin{aligned} \sum_{b \in B} q_b^{d*} z_b^d + \sum_{b \in B} q_b^{e*} z_b^e &= \sum_{b \in B} \mu_b (q_b^{d*} + q_b^{e*}) \\ &\leq \sum_{b \in B} \mu_b \mathbf{q}_F^{d*} \cdot \mathbf{x}_b \\ &= \sum_{f \in F} \kappa_f q_f^{d*} \end{aligned}$$

because attainability requires that $z_b^d = z_b^e = \mu_b$, for any $b = (\mathbf{x}, d) \in B$, and $\sum_{b \in B} \mu_b x_b = \kappa_f$, for any $f \in F$.

Similarly, the equilibrium optimality condition for firms implies that

$$q_f^{d*} + q_f^{e*} \leq 1,$$

for any $f \in F$. But this implies that

$$\begin{aligned} \mathbf{q}^* \cdot \mathbf{z} &= \sum_{b \in B} q_b^{d*} \mu_b + \sum_{b \in B} q_b^{e*} \mu_b + \sum_{f \in F} q_f^{e*} \kappa_f \\ &\leq \sum_{f \in F} q_f^{d*} \kappa_f + \sum_{f \in F} q_f^{e*} \kappa_f \\ &\leq \sum_{f \in F} \kappa_f = 1. \end{aligned}$$

This contradicts our initial hypothesis and proves that the equilibrium allocation must be efficient.

Proof of Proposition 2

Let C denote the set of attainable consumption vectors. Then the supporting hyperplane theorem tells us that there exists a non-negative price vector \mathbf{p}^* such that

$$\mathbf{p}^* \cdot \mathbf{c}^* = \sup \{ \mathbf{p}^* \cdot \mathbf{c} : \mathbf{c} \in C \}.$$

Without loss of generality, we normalize prices so that $\mathbf{p}^* \cdot \mathbf{c}^* = 1$. Now define the securities' prices \mathbf{q}^* as follows:

$$\begin{aligned} q_b^{d*} &= \mathbf{p}^* \cdot (\mathbf{a}_b^d, \mathbf{0}), \\ q_b^{e*} &= \mathbf{p}^* \cdot (\mathbf{0}, \mathbf{a}_b^e), \\ q_f^{e*} &= \mathbf{p}^* \cdot (\mathbf{0}, a_f^e), \text{ and} \\ q_f^{d*} &= 1 - q_f^{e*}. \end{aligned}$$

For any $(\mathbf{c}, \boldsymbol{\mu}, \boldsymbol{\kappa})$ satisfying (11a) and (12a),

$$\begin{aligned} 1 &= \mathbf{p}^* \cdot \mathbf{c}^* \\ &\geq \mathbf{p}^* \cdot \mathbf{c} \\ &= \mathbf{p}^* \cdot \left(\sum_{b \in B} \mu_b (\mathbf{a}_b^d, \mathbf{a}_b^e) + \sum_{f \in F} \kappa_f (\mathbf{0}, \mathbf{a}_f^e) \right) \\ &= \sum_{b \in B} \mu_b (q_b^{d*} + q_b^{e*}) + \sum_{f \in F} \kappa_f q_f^{d*} \\ &= \sum_{b \in B} \mu_b (q_b^{d*} + q_b^{e*}) + \sum_{b \in B} \mu_b \mathbf{x}_b \cdot \mathbf{q}_F^{d*}. \end{aligned}$$

It follows from this inequality that, for any bank b , with portfolio \mathbf{x} ,

$$\begin{aligned} q_b^{d*} + q_b^{e*} + \mathbf{x} \cdot \mathbf{q}_F^{d*} &\leq 1, \\ \implies q_b^{d*} + q_b^{e*} + \mathbf{x} \cdot (\mathbf{1} - \mathbf{q}_F^{e*}) &\leq 1, \\ \implies q_b^{d*} + q_b^{e*} - \mathbf{x} \cdot \mathbf{q}_F^{e*} &\leq 0. \end{aligned}$$

In other words, no bank can earn positive profits. But active firms must earn zero profits in equilibrium, because

$$\begin{aligned} 1 &= \mathbf{p}^* \cdot \mathbf{c}^* \\ &= \mathbf{p}^* \cdot \left(\sum_{b \in B} \mu_b^* (\mathbf{a}_b^d, \mathbf{a}_b^e) + \sum_{f \in F} \kappa_f^* (\mathbf{0}, \mathbf{a}_f^e) \right) \\ &= \sum_{b \in B} \mu_b^* (q_b^{d*} + q_b^{e*}) + \sum_{f \in F} \kappa_f^* q_f^{e*} \\ &= \sum_{b \in B} \mu_b^* (q_b^{d*} + q_b^{e*}) + 1 - \sum_{f \in F} \kappa_f^* q_f^{d*} \end{aligned}$$

so

$$\sum_{b \in B} \mu_b^* (q_b^{d^*} + q_b^{e^*}) - \sum_{f \in F} \kappa_f^* q_f^{d^*} = 0.$$

Then $\mu_b^* > 0$ implies $q_b^{d^*} + q_b^{e^*} - \mathbf{x} \cdot \mathbf{q}_F^{d^*} = 0$.

By definition, we have

$$q_f^{d^*} + q_f^{e^*} = 1,$$

so all firms are value maximizing.

Finally, the optimality of the representative consumer's choice follows from the fact that $U(\mathbf{c}) > U(\mathbf{c}^*)$ implies that $\mathbf{p}^* \cdot \mathbf{c} > \mathbf{p}^* \cdot \mathbf{c}^*$. Any portfolio \mathbf{z} that generates a consumption bundle \mathbf{c} must therefore have a value greater than 1 because we have defined security prices so that

$$\mathbf{q} \cdot \mathbf{z} = \mathbf{p} \cdot \mathbf{c}.$$

But this implies that \mathbf{z} does not belong to the budget set $\{\mathbf{z} : \mathbf{q} \cdot \mathbf{z} \leq 1\}$.

Proof of Proposition 3

The proof is by contradiction. Suppose, contrary to what we want to prove, that $\widehat{d} \geq \hat{\ell}$. Then the bank will be bankrupt with probability one and $d(a) = \lambda_b \ell(a)$ for all a . Hence the representative agent's utility is

$$\int_0^\infty \left\{ d(a) + \beta \int_{\ell/a}^\infty (\theta a - \ell) f(\theta) d\theta \right\} g(a) da = \int_0^\infty \left\{ \lambda_b \ell(a) + \beta \int_{\ell/a}^\infty (\theta a - \ell) f(\theta) d\theta \right\} g(a) da.$$

Consider the effect of reducing the face value of deposits so that \bar{d} satisfies

$$\frac{(1 - \beta) \ell}{2 - \beta - \lambda_b} < \bar{d} < \ell, \tag{17}$$

while holding ℓ constant at $\hat{\ell}$. Note that there exists a non-empty set of such values of \bar{d} because

$$\frac{(1 - \beta) \ell}{2 - \beta - \lambda_b} = \frac{(1 - \beta) \ell}{1 - \beta + 1 - \lambda_b} < \ell.$$

Then, for large values of a , we avoid the bankruptcy costs imposed on the banks. Let \bar{a} be the smallest value of a for which the banks are solvent: $\ell(\bar{a}) = \bar{d}$. Then the gain from reducing \bar{d} is

$$\int_{\bar{a}}^\infty \left\{ \beta (\ell(a) - \bar{d}) + \bar{d} - \lambda_b \ell(a) \right\} g(a) da = \int_{\bar{a}}^\infty \left\{ (\beta - 1) (\ell(a) - \bar{d}) + (1 - \lambda_b) \ell(a) \right\} g(a) da$$

which must be positive because the inequality (17) implies that the integrand of the right hand expression is positive:

$$\begin{aligned} (\beta - 1) (\ell(a) - \bar{d}) + (1 - \lambda_b) \ell(a) &> (\beta - 1) (\ell - \bar{d}) + (1 - \lambda_b) \bar{d} \\ &= (\beta - 1) \ell + (2 - \beta - \lambda_b) \bar{d} \\ &> (\beta - 1) \ell + (1 - \beta) \ell = 0. \end{aligned}$$

Proof of Proposition 5

Let $(\mathbf{c}^*, \mathbf{z}^*, \boldsymbol{\mu}^*, \boldsymbol{\kappa}^*)$ be an equilibrium and let $(d, \mathbf{x}_b) = b \in B$ be a fixed but arbitrary type of bank such that $\mu_b^* > 0$. For any firm-type $f = (\ell, j) \in F$ such that $x_{bf} > 0$, recall that the loan repayment in state s is,

$$a_{fs}^d = \begin{cases} \ell & \text{if } \ell \leq A_{js} \\ \lambda_f A_{js} & \text{if } \ell > A_{js} \end{cases}$$

for each state s . Then the total amount owed to bank b is

$$\ell_b = \sum_{f=(\ell,j) \in F} x_{bf} \ell$$

and the actual amount received in state s is

$$\ell_{bs} = \sum_{f \in F} x_{bf} a_{fs}^d$$

for each state s . The equity of a bank of type b pays out $\max\{\ell_{bs} - d, 0\}$ units of consumption to shareholders in any state s and the holders of deposits receive

$$d_{bs} = \begin{cases} d & \text{if } d \leq \ell_{bs} \\ \lambda_b \ell_{bs} & \text{if } d > \ell_{bs} \end{cases}$$

for any state s .

In order to prove the proposition, we need to assume that the sets D and L are sufficiently rich.

Definition 11 *The sets D and L are sufficiently rich in the following sense: for any (j, s) , $A_{js} \in L$ and, for any (b, s) , $\ell_{bs} \in D$.*

The first requirement ensures that any producer can choose the face value of his debt to coincide with the output of his chosen technology in a particular state. The second requirement ensures that a banker can choose the amount of deposits to coincide with the revenue

from a portfolio of loans in a particular state. Note that these quantities are determined by the respective types of firm and bank.

Step 1: *The value of bank b 's equity is positive if and only if $\ell_b > d$.*

The dividends paid by the bank are non-negative in each state because of limited liability. Then it is clear that the value of equity is strictly positive if and only if $\ell_{bs} > d$ for at least one state s . Because default is costly, an active firm $f = (\ell, j)$ will choose a debt level $\ell_f \leq A_{fs}$ and an active bank $b = (d, \mathbf{x}_b)$ will choose a level of deposits $d \leq \sum_{f \in F} x_{bf} \ell_f$. We claim that the value of bank b 's equity is positive if and only if $d < \ell_b$. To see this, note first that, if $\ell_b = d$, the payment to equity holders (as defined above) is zero in each state and, second, if $\ell_b > d$, the payment must be positive in state S at least, because in that state each firm f with $x_{bf} > 0$ repays the face value of its debt in full, the bank receives ℓ_b , and the equity holders receive $\ell_b - d > 0$.

Step 2: *For each f such that $x_{bf} > 0$, there exists a state s_f such that firm f is solvent if and only if $s \geq s_f$. Similarly, there exists a state s_b such that bank b is solvent if and only if $s \geq s_b$.*

A firm $f = (\ell, j)$ is insolvent in state s if and only if $\ell > A_{js}$. Let s_f be the smallest state such that f is solvent. There must be such a state because the firm is solvent at least in state S . The fact that A_{js} is increasing in s implies that f is solvent if and only if $s \geq s_f$.

Similarly, we can show that there is a state s_b such that bank b is solvent if and only if $s \geq s_b$. Let s_b denote the smallest state in which the bank is solvent. There must be such a state because the bank is solvent at least in state S . It is clear that ℓ_{bs} is non-decreasing in s because A_{js} is increasing in s for every j and this implies that ℓ_{fs} and ℓ_{bs} are non-decreasing in s , for every f . Then it is clear that d_{bs} is non-decreasing in s because ℓ_{bs} is non-decreasing in s . From this observation it follows that the bank is solvent if and only if $s \geq s_b$.

Step 3: *The face value of deposits satisfies $d = \ell_{bs_b}$.*

To prove this claim, we have to consider two cases. First, suppose that $s_b = 1$, that is bank b is solvent in every state, and $d < \ell_{b1}$. Then increasing d to $d' = \ell_{b1}$, say, will increase the payout to depositors by in every state and reduce the payout to equity holders by the same amount. This increases effective consumption by $(1 - \beta)(d' - d)$ in each state, contradicting the constrained efficiency of the equilibrium.

The second possibility is that $s_b > 1$ and $d < \ell_{bs_b}$. Suppose the bank increases deposits to $d' = \ell_{bs_b}$. This will not have any effect in states $s < s_b$ because the bank is in default, depositors are receiving $\ell_{bs} < d$, and bank equity holders are receiving nothing. In states $s \geq s_b$, on the other hand, the net effect will be an increase in effective consumption of $(1 - \beta)(d' - d)$. This again contradicts constrained efficiency.

Step 4: *Bank b 's equity has no value: $d = \ell_b$.*

Suppose to the contrary that $d = \ell_{bs_b} < \ell_b$. Then there must exist at least one firm that is bankrupt in state s_b . Otherwise, the firms' repayment would be $\ell_{bs_b} = \ell_b$. Suppose that a firm-type $f = (\ell, j)$ with $x_{bf} > 0$ is bankrupt in state s_b and consider the effect of reducing the borrowing of f to $\ell'_f = A_{js_b}$. This change has no impact on the viability of the bank in state s_b because firm f is now paying A_{js_b} in state s_b , instead of $\lambda_f A_{js_b}$. So the bank is solvent in states $s \geq s_b$ as before. Also, the change does have an effect on the solvency of the firm, as it is now solvent in state s_b and hence in all states $s \geq s_b$.

Note that none of the payoffs to the debt and equity of the bank or the firm change in states $s < s_b$. Moreover, there are no changes to the payoff to bank deposits in states $s \geq s_b$ because the bank is solvent in all these states and hence pays d to depositors. The changes in payoff affect only the returns to equity in the states $s \geq s_b$. Consider first the bank's equity. In states $s \geq s_b$, the payoff to the bank's equity will increase because of the increase in the firm's repayment. In fact, the change in the payoff of the bank's equity is

$$\Delta_{b,s} = \begin{cases} A_{js_b} - \lambda_f A_{js} & \text{if } s_b \leq s < s_f \\ A_{js_b} - \ell_f & \text{if } s_f \leq s \end{cases},$$

because the firm will pay A_{js_b} in each state and was previously paying $\lambda_f A_{js}$ in states $s_b \leq s < s_f$ and ℓ_f in states $s_f \leq s$.

The return to the equity of firm f is increased by $A_{js} - A_{js_b}$ for all $s \geq s_b$, so the change in the firm's equity is

$$\Delta_{f,s} = \begin{cases} A_{js} - A_{js_b} & \text{if } s_b \leq s < s_f \\ \ell - A_{js_b} & \text{if } s_f \leq s \end{cases}.$$

The transfer between bank equity holders and firm equity holders has no effect on total consumption. The net increase in consumption is the sum of $\Delta_{b,s}$ and $\Delta_{f,s}$, that is,

$$\begin{aligned} \Delta_{b,s} + \Delta_{f,s} &= \begin{cases} A_{fs_b} - \lambda_f A_{fs} + A_{js} - A_{js_b} & \text{if } s_b \leq s < s_f \\ A_{js_b} - \ell_f + \ell_f - A_{js_b} & \text{if } s_f \leq s \end{cases} \\ &= \begin{cases} A_{js} - \lambda_f A_{js} & \text{if } s_b \leq s < s_f \\ 0 & \text{if } s_f \leq s \end{cases}. \end{aligned}$$

Thus, the net gain for equity holders is the saving in default costs $(1 - \lambda_f) A_{js}$ in the states $s_b \leq s < s_f$. The possibility of such a gain contradicts the constrained efficiency of the equilibrium.

This completes the proof that bank equity has no value.

Technology shocks in (16) and the single factor model

To compare the specification in (16) with the single factor model, note that we can represent the productivity of technology j as $A_{js} = \theta_j A$, where A is an aggregate shock and θ_j is an idiosyncratic shock, with marginal distributions given by

$$A = \begin{cases} a_H & \text{w. prob. } \delta \\ a_M & \text{w. prob. } 1 - \delta \end{cases}$$

and

$$\theta_j = \begin{cases} 1 & \text{w. prob. } \delta \\ 1 & \text{w. prob. } 1 - \delta - (n-1)\varepsilon/n \\ a_L/a_M & \text{w. prob. } (1 - \delta - \varepsilon)/n \end{cases} .$$

In the single factor model of Section 4.1, we assumed that the shock θ_j was i.i.d. across technologies and independent of A . In (16) we assume that the idiosyncratic shock θ_j is degenerate and equal to one when $A = a_H$, while it takes the values 1 and a_L/a_M when $A = a_M$. In addition, we assume the idiosyncratic shocks are negatively correlated across technologies, in the sense that the realization of θ_j is equal to $a_L/a_M < 1$ for at most one technology j in any state.

Proof of Proposition 7.

Banks holding risk free debt In the case of a bank holding only the debt of firms with $\ell_j = a_L$, the portfolio \mathbf{x} is indeterminate subject to the constraint $\sum_{j=1}^n x_j^L = 1$. The optimal capital structure for the bank is to issue the maximum amount of deposits, $d = a_L$. The expected utility generated by the bank and the firms whose debt it holds will be

$$\begin{aligned} & d + \varepsilon\beta(a_H - d) + \delta\beta(a_M - d) + (1 - \delta - \varepsilon)\beta\left(\frac{n-1}{n}a_M + \frac{1}{n}a_L - d\right) \\ = & a_L + \varepsilon\beta(a_H - a_L) + \delta\beta(a_M - a_L) + (1 - \delta - \varepsilon)\beta\left(\frac{n-1}{n}a_M - \frac{n-1}{n}a_L\right) \\ = & a_L + \varepsilon\beta(a_H - a_L) + \left(\delta + (1 - \delta - \varepsilon)\frac{n-1}{n}\right)\beta(a_M - a_L). \end{aligned}$$

Banks holding safe and risky debt Now suppose that the bank lends to a mixture of safe and risky firms. We can focus here on the loans made by the bank to firms with technology j but possibly different capital structures. Suppose that γ units of capital are invested in safe firms with $\ell_j = a_L$ and $1 - \gamma$ units of capital are invested in risky or very risky firms, that is, firms that have a capital structure $\ell_j \in \{a_H, a_M\}$. There is no need to distinguish safe firms according to the technology they use: from the point of view of banks and shareholders, who hold their debt and equity, they are identical. Let x_j^H and x_j^M denote the fraction of $1 - \gamma$ invested in firms with technology j and ℓ_j equal to a_H and a_M ,

respectively. Note that $\sum_{j=1}^n (x_j^H + x_j^M) = 1$ and that the amounts invested in firms with technology j and ℓ_j equal to a_H and a_M are $x_j^H (1 - \gamma)$ and $x_j^M (1 - \gamma)$, respectively.

Suppose the bank chooses a level of deposits d . Since the safe banks pay a_L for sure, the bank will fail if and only if payment from the risky banks is less than $d - \gamma a_L$. Now suppose that we split the bank into two banks, one of which funds safe firms and the other funds risky firms. The safe bank invests one unit in safe firms and issues deposits $d^S = a_L$ and the risky bank invests one unit in a portfolio $\{x_j^H, x_j^M\}_{j=1}^n$ of risky firms and issues deposits d^R . The expected utility from the safe bank is denoted by U^S and the expected utility from the risky bank is denoted by U^R . What is the difference between these two banks and the combined bank we started with? Note that the risky bank will default if and only if the unified bank defaults with a positive probability. If there is no probability of default, there is no difference in the expected utility generated by the two structures. On the other hand, if there is a positive probability of default, the separated banks will generate a higher expected utility, because the safe bank does not default whereas the combined bank does default in some states. In fact, the gain in expected utility by separating the banks is precisely the probability of default multiplied by the default cost $(1 - \lambda_b) d_b^S$.

Thus, either there is no gain from mixing safe and risky debt in the banks portfolio or, if the mixture of safe and risky debt causes the bank to default with positive probability, there is a loss.

Safe banks holding risky debt Now suppose that a bank chooses a portfolio $\mathbf{x} = \{(x_j^H, x_j^M, 0)\}_{j=1}^n$ where x_j^H is the measure of firms of type j with $\ell_j = a_H$, x_j^M is the measure of firms of type j with $\ell_j = a_M$, and we assume that no firms with $\ell_j = a_L$ are included. The portfolio \mathbf{x} has no impact in states $s = n + 1, n + 2$ because all technologies have identical payoffs in these states. Now consider the states $s = 1, \dots, n$ and let ρ_j denote the repayment of all firms when type j has productivity a_L . Then

$$\rho_j = \sum_{i \neq j} (x_i^H \lambda_f + x_i^M) a_M + (x_j^H + x_j^M) \lambda_f a_L.$$

Without essential loss of generality, we can order the types of firms so that $\rho_j \leq \rho_{j+1}$ for $j = 1, \dots, n - 1$. The bank wants to maximize the face value of deposits subject to the no-default constraint $d \leq \rho_1$. To do that, it must choose a portfolio \mathbf{x} such that $(x_j^H, x_j^M) = (0, \frac{1}{n})$ for all $j = 1, \dots, n$, that is, a simple portfolio. Having done so, the value of deposits it can safely issue is

$$d = \frac{n-1}{n} a_M + \frac{1}{n} \lambda_f a_L.$$

The expected utility generated by the bank and the firms whose debt it holds will be

$$\begin{aligned} & d + \beta(1 - \delta - \varepsilon) \left(\frac{n-1}{n} a_M + \frac{1}{n} \lambda_f a_L - d \right) + \beta\delta(a_M - d) + \beta\varepsilon(a_H - d) \\ = & d + \beta\delta(a_M - d) + \beta\varepsilon(a_H - d). \end{aligned}$$

Risky banks holding risky debt We split the analysis in two parts, considering first that case where $d \leq a_M$ and, second, the case where $d > a_M$.

i) Suppose that there is a positive probability of the bank defaulting, but that $d \leq a_M$. This means that default only occurs in states $s = 1, \dots, n$. As before, the portfolio \mathbf{x} is irrelevant in states $s = n+1, n+2$ so we restrict attention to the states $s = 1, \dots, n$. With our usual convention that $\rho_j \leq \rho_{j+1}$, there exists a technology k such that, $d > \rho_j$ for $j = 1, \dots, k$ and $d \leq \rho_j$ for $j = k+1, \dots, n$. (The bank will never choose to default with probability one). The expected utility of the bank's depositors and shareholders will be

$$\begin{aligned} \frac{1}{n} \lambda_b \sum_{j=1}^k \rho_j + \frac{n-k}{n} d + \frac{1}{n} \beta \sum_{j=k+1}^n (\rho_j - d) & \leq \frac{1}{n} \lambda_b \sum_{j=1}^k \rho_j + \frac{n-k}{n} d + \frac{1}{n} \sum_{j=k+1}^n (\rho_j - d) \\ & = \frac{1}{n} \lambda_b \sum_{j=1}^k \rho_j + \frac{1}{n} \sum_{j=k+1}^n \rho_j, \end{aligned}$$

because $\beta < 1$. Now

$$\begin{aligned} \frac{1}{n} \lambda_b \sum_{j=1}^k \rho_j & = \frac{1}{n} \lambda_b \sum_{j=1}^k \left(\sum_{i \neq j} (x_i^H \lambda_f + x_i^M) a_M + (x_j^H + x_j^M) \lambda_f a_L \right) \\ & = \frac{1}{n} \lambda_b \left(\sum_{j=k+1}^n (x_j^H \lambda_f + x_j^M) a_M + \sum_{j=1}^k (x_j^H + x_j^M) \lambda_f a_L \right) \\ & \quad + \frac{1}{n} \lambda_b \sum_{j=2}^k (x_j^H \lambda_f + x_j^M) a_M - \frac{1}{n} \lambda_b \sum_{j=2}^k (x_j^H + x_j^M) \lambda_f a_L \\ & \quad + \frac{1}{n} \lambda_b \sum_{j=2}^k \left(\sum_{i \neq j} (x_i^H \lambda_f + x_i^M) a_M + (x_j^H + x_j^M) \lambda_f a_L \right) \end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{n} \lambda_b \left(\sum_{j=k+1}^n (x_j^H \lambda_f + x_j^M) a_M + \sum_{j=1}^k (x_j^H + x_j^M) \lambda_f a_L \right) \\
&\quad + \frac{1}{n} \sum_{j=2}^k (x_j^H \lambda_f + x_j^M) a_M - \frac{1}{n} \sum_{j=2}^k (x_j^H + x_j^M) \lambda_f a_L \\
&\quad + \frac{1}{n} \sum_{j=2}^k \left(\sum_{i \neq j} (x_i^H \lambda_f + x_i^M) a_M + (x_j^H + x_j^M) \lambda_f a_L \right),
\end{aligned}$$

because $\lambda_b < 1$ and $(x_j^H \lambda_f + x_j^M) a_M \geq (x_j^H + x_j^M) \lambda_f a_L$,

$$\begin{aligned}
&= \frac{1}{n} \lambda_b \left(\sum_{j=k+1}^n (x_j^H \lambda_f + x_j^M) a_M + \sum_{j=1}^k (x_j^H + x_j^M) \lambda_f a_L \right) \\
&\quad + \frac{1}{n} \sum_{j=2}^k \left(\sum_{i \neq j} (x_i^H \lambda_f + x_i^M) a_M + (x_j^H \lambda_f + x_j^M) a_M \right) \\
&= \frac{1}{n} \lambda_b \left(\sum_{j=k+1}^n (x_j^H \lambda_f + x_j^M) a_M + \sum_{j=1}^k (x_j^H + x_j^M) \lambda_f a_L \right) \\
&\quad + \frac{k-1}{n} \sum_{j=1}^n (x_j^H \lambda_f + x_j^M) a_M.
\end{aligned}$$

Substituting this upper bound into the expression for expected utility, we obtain the inequality

$$\begin{aligned}
&\frac{1}{n} \lambda_b \sum_{j=1}^k \rho_j + \frac{1}{n} \sum_{j=k+1}^n \rho_j \\
&\leq \frac{1}{n} \lambda_b \left(\sum_{j=k+1}^n (x_j^H \lambda_f + x_j^M) a_M + \sum_{j=1}^k (x_j^H + x_j^M) \lambda_f a_L \right) + \frac{k-1}{n} \sum_{j=1}^n (x_j^H \lambda_f + x_j^M) a_M \\
&\quad + \frac{1}{n} \sum_{j=k+1}^n \left(\sum_{i \neq j} (x_i^H \lambda_f + x_i^M) a_M + (x_j^H + x_j^M) \lambda_f a_L \right)
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{n} \lambda_b \left(\sum_{j=k+1}^n (x_j^H \lambda_f + x_j^M) a_M + \sum_{j=1}^k (x_j^H + x_j^M) \lambda_f a_L \right) + (1 - \lambda_b) \frac{1}{n} \sum_{j=k+1}^n (x_j^H \lambda_f + x_j^M) a_M \\
&- (1 - \lambda_b) \frac{1}{n} \sum_{j=k+1}^n (x_j^H + x_j^M) \lambda_f a_L + \frac{k-1}{n} \sum_{j=1}^n (x_j^H \lambda_f + x_j^M) a_M \\
&+ \frac{1}{n} \sum_{j=k+1}^n \left(\sum_{i \neq j} (x_i^H \lambda_f + x_i^M) a_M + (x_j^H + x_j^M) \lambda_f a_L \right) \\
&= \frac{1}{n} \lambda_b \left(\sum_{j=k+1}^n (x_j^H + x_j^M) \lambda_f a_L + \sum_{j=1}^k (x_j^H + x_j^M) \lambda_f a_L \right) + \frac{k-1}{n} \sum_{j=1}^n (x_j^H \lambda_f + x_j^M) a_M \\
&+ \frac{1}{n} \sum_{j=k+1}^n \left(\sum_{i \neq j} (x_i^H \lambda_f + x_i^M) a_M + (x_j^H \lambda_f + x_j^M) a_M \right)
\end{aligned}$$

because $\lambda_b < 1$ and $\sum_{j=k+1}^n (x_j^H + x_j^M) \lambda_f a_L < \sum_{j=k+1}^n (x_j^H \lambda_f + x_j^M) a_M$. So

$$\begin{aligned}
&\frac{1}{n} \lambda_b \sum_{j=1}^k \rho_j + \frac{1}{n} \sum_{j=k+1}^n \rho_j \\
&\leq \frac{1}{n} \lambda_b \left(\sum_{j=k+1}^n (x_j^H + x_j^M) \lambda_f a_L + \sum_{j=1}^k (x_j^H + x_j^M) \lambda_f a_L \right) \\
&+ \frac{k-1}{n} \sum_{j=1}^n (x_j^H \lambda_f + x_j^M) a_M + \frac{1}{n} \sum_{j=k+1}^n \left(\sum_{i \neq j} (x_i^H \lambda_f + x_i^M) a_M + (x_j^H \lambda_f + x_j^M) a_M \right) \\
&\leq \frac{1}{n} \lambda_b \lambda_f a_L + \frac{k-1}{n} a_M + \frac{n-k}{n} a_M = \frac{1}{n} \lambda_b \lambda_f a_L + \frac{n-1}{n} a_M,
\end{aligned}$$

because $\sum_{j=1}^n (x_j^H \lambda_f + x_j^M) \leq \sum_{j=1}^n (x_j^H + x_j^M) = 1$. But the last expression in this series of inequalities is the representative consumer's expected utility when the bank lends only to firms that use a single technology j and choose the capital structure $\ell = a_M$ and the level of deposits is $d = a_M$. It is easy to check that for any other portfolio and face value of deposits one of the inequalities is strict, so this is the unique policy that maximizes expected utility in the states $s = 1, \dots, n$ when $d \leq a_M$ and the probability of default is positive.

Now let us check that this policy is optimal in the states $s = n+1, n+2$. The expected utility in states $s = n+1, n+2$ is

$$\frac{\delta}{\delta + \varepsilon} a_M + \frac{\varepsilon}{\delta + \varepsilon} (a_M + \beta (a_H - a_M)).$$

For any other portfolio $\{(x_j^H, x_j^M, 0)\}_{j=1}^n$, the bank will be in default in state $s = n+1$ if

$\sum_{j=1}^n x_j^H > 0$ and the expected utility in states $s = n + 1, n + 2$ is

$$\frac{\delta}{\delta + \varepsilon} \lambda_b \left(\sum_{j=1}^n (x_j^H \lambda_f + x_j^M) a_M \right) + \frac{\varepsilon}{\delta + \varepsilon} (a_M + \beta (a_H - a_M)) < \\ \frac{\delta}{\delta + \varepsilon} a_M + \frac{\varepsilon}{\delta + \varepsilon} (a_M + \beta (a_H - a_M)).$$

If $\sum_{j=1}^n x_j^H = 0$, the bank is not in default in either state and the payoff is

$$\frac{\delta}{\delta + \varepsilon} \left(\sum_{j=1}^n x_j^M a_M \right) + \frac{\varepsilon}{\delta + \varepsilon} (a_M + \beta (a_H - a_M)) = \\ \frac{\delta}{\delta + \varepsilon} a_M + \frac{\varepsilon}{\delta + \varepsilon} (a_M + \beta (a_H - a_M))$$

Thus, the unique optimal policy is to set $d = a_M$ and $x_j^M = 1$ for some j as long as $d \leq a_M$.

ii) Now consider the case in which $d > a_M$. In that case, the bank always defaults states $s = 1, \dots, n + 1$. In states $s = 1, \dots, n$, the expected utility will be

$$\frac{1}{n} \sum_{j=1}^n \lambda_b \left(\sum_{i \neq j} (x_i^H \lambda_f + x_i^M) a_M + (x_j^H + x_j^M) \lambda_f a_L \right)$$

and in states $s = n + 1, n + 2$ it will be

$$\frac{\delta}{\delta + \varepsilon} \lambda_b \sum_{j=1}^n (x_j^H \lambda_f + x_j^M) a_M + \frac{\varepsilon}{\delta + \varepsilon} (d + \beta (a_H - d)).$$

The choice of d will be the maximum that allows the bank to remain solvent in state $s = n + 2$, that is, $d = \sum_{j=1}^n (x_j^H a_H + x_j^M a_M)$. Letting $x^H = \sum_{j=1}^n x_j^H$ and $x^M = \sum_{j=1}^n x_j^M$, we can rewrite the expected utility as

$$\frac{\delta}{\delta + \varepsilon} \lambda_b (x^H \lambda_f + x^M) a_M + \frac{\varepsilon}{\delta + \varepsilon} (d + \beta (a_H - d))$$

in states $s = n + 1, n + 2$ and

$$\lambda_b \left(\frac{n-1}{n} (x^H \lambda_f + x^M) a_M + \frac{1}{n} \lambda_f a_M \right)$$

in states $s = 1, \dots, n$. This expression is linear in (x^H, x^M) , so at least one of the extreme points $(x^H, x^M) = (0, 1)$ or $(x^H, x^M) = (1, 0)$ must be an optimum. Since we assume that $d > a_M$, this case is only observed if $x^H = 1$.

Proof of Proposition 8.

As shown in the text, if the representative bank lends to safe firms ($\ell = a_L$) the expected utility is

$$a_L + \frac{n-1}{n}\beta(a_M - a_L).$$

If the representative bank issues safe debt (deposits) and lends to risky firms ($\ell = a_M$), on the other hand, the expected utility is

$$\frac{n-1}{n}a_M + \frac{1}{n}\lambda_f a_L.$$

It is strictly optimal to issue safe debt and lend to risky firms if

$$a_L + \frac{n-1}{n}\beta(a_M - a_L) < \frac{n-1}{n}a_M + \frac{1}{n}\lambda_f a_L,$$

which is equivalent to

$$\frac{1}{1-\beta} \left(\frac{n-\lambda_f}{n-1} - \beta \right) a_L < a_M.$$

Proof of Proposition 9

The argument in the proof of Proposition 7 left us with the following candidates for an optimal bank policy.

1. The bank invests in firms with $\ell = a_L$ and $d = a_L$. The firms' types are irrelevant because firm debt is risk free. The expected utility in equilibrium is

$$a_L + \beta \left((1 - \delta - \varepsilon) \left(\frac{n-1}{n}a_M + \frac{1}{n}a_L \right) \right) + \delta a_M + \varepsilon a_H - a_L.$$

- 2a. The bank invests in firms with $\ell = a_M$. The portfolio is defined by $x_j^M = \frac{1}{n}$ for all j and the face value of deposits is $d = \frac{n-1}{n}a_M + \frac{1}{n}\lambda_f a_L$. The expected utility in equilibrium is

$$\frac{n-1}{n}a_M + \frac{1}{n}\lambda_f a_L + \beta(\delta + \varepsilon) \left(\frac{\delta}{\delta + \varepsilon}a_M + \frac{\varepsilon}{\delta + \varepsilon}a_H - \frac{n-1}{n}a_M - \frac{1}{n}\lambda_f a_L \right)$$

- 2b. The bank invests in firms with $\ell = a_M$. The portfolio is defined by $x_j^M = 1$ for some j and the face value of deposits is $d = a_M$. The expected utility in equilibrium is

$$(1 - \delta - \varepsilon) \left(\frac{n-1}{n}a_M + \frac{1}{n}\lambda_b \lambda_f a_L \right) + (\delta + \varepsilon) a_M + \varepsilon \beta (a_H - a_M).$$

3. The bank invests in firms with $\ell = a_H$. The portfolio is defined by $x^H = 1$ (the distribution over j is irrelevant) and the face value of deposits is $d = a_H$. The expected utility in equilibrium is

$$\varepsilon a_H + \lambda_b \left(\delta \lambda_f a_M + (1 - \delta - \varepsilon) \left(\frac{n-1}{n} \lambda_f a_M + \frac{1}{n} \lambda_f a_L \right) \right).$$

Suppose that δ and ε converge to zero. The expected utilities in the different cases converge to

$$a_L + \beta \left(\frac{n-1}{n} a_M - \frac{n-1}{n} a_L \right), \quad (\text{Case 1})$$

$$\frac{n-1}{n} a_M + \frac{1}{n} \lambda_f a_L, \quad (\text{Case 2a})$$

$$\left(\frac{n-1}{n} a_M + \frac{1}{n} \lambda_b \lambda_f a_L \right), \quad (\text{Case 2b})$$

and

$$\lambda_b \lambda_f \left(\frac{n-1}{n} a_M + \frac{1}{n} a_L \right), \quad (\text{Case 3})$$

respectively. Finally, Proposition 9 guarantees that Case 2a dominates Case 1. Thus, Case 2a dominates all other cases for values of δ and ε sufficiently close to zero.