

Equilibrium Theory of Banks' Capital Structure*

Douglas Gale
New York University

Piero Gottardi
European University Institute

February 6, 2017

Abstract

We study an environment where the capital structure of banks and firms are jointly determined in equilibrium, so as to balance the benefits of the provision of liquidity services by bank deposits with the costs of bankruptcy. The risk in the assets held by firms and banks is determined by the technology choices by firms and the portfolio diversification choices by banks. We show competitive equilibria are efficient and the equilibrium level of leverage in banks and firms depend on the nature of the shocks affecting firm productivities. When these shocks are co-monotonic, banks optimally choose a zero level of equity. Thus all equity should be in firms, where it does “double duty,” protecting both firms and banks from default. On the other hand, if productivity shocks have an idiosyncratic component, portfolio diversification by banks may be a more effective buffer against these shocks and, in these cases, it may be optimal for banks, as well as firms, to issue equity.

1 Introduction

The Financial Crisis of 2007-08 started a vigorous debate about the regulation of the banking system, much of it focused on bank capital regulation. It is widely accepted that capital adequacy regulation in the pre-crisis period did not ensure that banks had sufficient capital to weather the storm. The risk weights used under Basel II underestimated the risk of certain asset classes and regulatory arbitrage further reduced its effectiveness. In response, regulators at both the national and international level have sought to introduce more stringent capital adequacy requirements.

Although new policies are already being put in place, our theoretical understanding of the role of bank capital lags behind. What is the market failure that requires capital regulation? What role do capital requirements play in the corporate governance of large and complex

*We are grateful to Franklin Allen and Cyril Monnet for very helpful comments. We thank the participants at seminars and conferences at New York University, Cambridge University, the University of Essex, The House of Finance in Stockholm, and the Copenhagen Business School for their comments and questions.

banks? What are the costs of regulation, directly in the form of compliance costs and indirectly in the form of distortions of economic decisions?

Capital regulation has its roots in bank supervision, which traditionally focuses on the “safety and soundness” of individual banks. Not surprisingly, recent proposals for increasing the stability of the financial system are simply stronger versions of the policies aimed at insuring the resilience of individual banks. Similarly, much of the theoretical literature on capital regulation focuses on the behavior of individual banks, rather than the financial system. This is unfortunate, because macroprudential regulation is concerned with the stability of the financial system as a whole. It requires an understanding of systemic risk, as distinct from the risk of individual banks. In addition, capital regulation may have macroeconomic effects. For example, a global increase in bank capital will have an effect on the cost of capital. Finally, stability is not the only objective of financial regulation. Its objectives should include an efficient, innovative, and competitive financial system. This requires an understanding of the welfare economics of regulation.

For all these reasons, we take the view that a general equilibrium approach is needed. A useful starting point is to identify the conditions under which *laissez-faire* equilibrium is efficient, as a precursor to identifying the market failures that makes regulation necessary and the distortions that may be introduced by regulation.

Admati and Hellwig (2013) have argued that the seminal Modigliani and Miller (1958) paper on capital structure should be the starting point for any discussion of capital regulation. We agree that Modigliani and Miller (1958) is the cornerstone of the literature on corporate capital structure, but an explanation of the determinants of bank capital structure cannot be restricted to the environment considered by Modigliani and Miller. In analyzing the capital structure of banks, it is important to recognize their role as intermediaries, as well as the fact that bank deposits are not simply debt claims, but are also valued by agents because they function as *money*. In this paper, we assume that deposits are used for transactions and this function gives rise to a spread between the returns on equity and deposits. At the same time, the higher a bank’s debt, the higher the probability and hence the cost of its default.¹

Unlike the Modigliani-Miller model, in which capital structure is irrelevant, in our model capital structure matters. A bank’s equilibrium capital structure is determined by a trade-off between the funding advantage of deposits and the possibility of costly default². In addition, the role of banks as intermediaries creates a link between the capital structure of banks and the capital structure of the firms that borrow from the banks. The higher firms’ leverage, the riskier their debt, and the higher the probability of bank default, other things being equal. This interdependence between the capital structures in the corporate sector and the banking sector is a major theme of this paper.

In this paper we study a representative agent economy consisting of consumers, firms

¹We assume that when a bank fails, because it is unable to meet the demand for withdrawals, a fraction of the value of its assets is lost.

²A similar trade-off arises in models of optimal capital structure where the firm balances the tax advantages of debt and the costs of default.

and banks. Consumers have an initial endowment of capital goods and want consumption goods. Firms have access to a variety of risky technologies that use capital goods to produce consumption goods. Firms issue equity to households and borrow from banks in order to fund the purchase of capital goods. Banks lend to firms and issue equity and deposits to households to fund their loan portfolios. Households purchase equity in firms and make deposits in banks to fund their future consumption. Only banks can lend to firms and 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 and 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 firms and banks can make different choices regarding technologies (in the case of firms), loan portfolios (in the case of banks), and capital structures (in the case of both firms and banks). These choices affect the risk characteristics of the securities issued by banks and firms and result in a large array of potential securities being priced and traded. We assume markets for these potential securities are competitive and *complete*, in the sense that there is a market and a market-clearing price for each type of security that could be issued.

In this framework, we obtain analogues of the fundamental theorems of welfare economics. First, we show that a competitive equilibrium, where firms and banks maximize their market value, is constrained efficient.³ Second, we show that any constrained efficient allocation can be decentralized⁴ as an equilibrium.⁵ Although these results are counterparts of the usual welfare theorems, the assumptions are demanding. In this framework, both firms and banks are innovating by creating new securities. When a firm changes its capital structure, it produces new forms of risky debt and risky equity. Similarly, when a bank changes its capital structure and portfolio, it produces new forms of risky deposits and risky equity. The coordination of these activities depends on the existence of markets for securities that do not exist in equilibrium.

We then turn to the critical question: *What is the equilibrium, constrained efficient level of equity for banks and firms?* The welfare theorems are crucial for the analysis of this question. Solving for an equilibrium is difficult, but since we know the equilibrium allocation is constrained efficient, we can use the necessary conditions for efficiency to characterize the equilibrium allocation. Our first result shows that, if the technologies available to firms satisfy a property known as *co-monotonicity*, which implies that their productivities are

³The first best or a Pareto efficient allocation cannot be attained because of the restriction to debt and equity as the funding instruments available to firms and banks. However, if we similarly restrict the planner to the allocations attainable using debt and equity, using so a notion of constrained efficiency, he would not be able to improve on the *laissez faire* equilibrium.

⁴Since there is a representative consumer lump sum taxes and transfers are not needed to decentralize the constrained efficient allocation.

⁵In this paper we ignore asymmetric information in order to focus on the welfare properties of the choices made by banks and firms in a basic competitive environment. There is a large and well known literature on moral hazard and risk shifting in banks. These considerations, as well as the possibility of bank bailouts in the event of default may distort the choice of the capital structure and introduce inefficiencies.

positively correlated, the value of bank capital will be zero in equilibrium. This is a striking result, but one that has an intuitive explanation. The fundamental source of uncertainty in the economy is the randomness of firms' productivity. When a firm receives a negative shock to productivity, it may be forced to default on its bank loans. This in turn makes bank loans risky and may trigger bank default as well. Equity in firms and banks represents a buffer that can absorb (some) losses and protect against costly default. The central question is the efficient allocation of equity between the corporate and banking sectors. In this regard, the equity in the firms' capital structure does "double duty." By protecting the firms against default, it also shields the banks from default. In the special case of co-monotonic technologies, this last consideration proves to be key. No matter what are the relative costs of default in banks and firms, it is always efficient to put all the equity in the corporate sector, where shocks are first felt.

The case of co-monotonic technologies is however special. If the productivity shocks have an idiosyncratic component banks may reduce their risk exposure by lending to firms using different technologies. In this case, the probability that any given technology receives a negative productivity shock is smaller than the probability that some of the available technologies receive a negative shock. Hence a diversified bank has a comparative advantage in providing a buffer against these shocks. First, it only needs to hold a small amount of equity because, in a typical state, only a small fraction of its portfolio of loans defaults. Second, because there is often some type of firm that is defaulting, the bank's equity buffer is needed more often than the firms' equity buffer. In such cases, it may be efficient for banks, rather than firms, to issue equity. In general, both banks and firms may find it (privately and socially) optimal to 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 ?? examines the properties of banks' capital structure in equilibrium. First, Section ?? shows that when the technologies are co-monotonic, bank equity has zero value in equilibrium. Then Section 4.2 characterizes the equilibrium capital structures in an environment where productivity shocks have an idiosyncratic component, finding that they feature a positive value of equity. Section 5 concludes. Proofs are collected in the appendix.

1.1 Related literature

The classic paper of Modigliani and Miller (1958) provides a benchmark in which 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, term structure, seniority and incentive problems, in the choice of corporate capital structure. A (non-representative) sample of this literature includes 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) and Hackbarth and Mauer (2012). In several cases the optimal capital structure is shown to be determinate.

As we have already noted, the key ingredients of our model of bank capital structure are:

- the interdependence between banks’ and firms’ capital structure, due to the intermediation role of banks;
- the fact that deposits earn a liquidity premium because of their use in transactions;
- the presence of costly default.

On the role of bank deposits as a source of funding for banks, Diamond and Dybvig (1983) and Diamond (1984) show they constitute the optimal form of funding that provide liquidity insurance to depositors or delegated monitoring for investors. Gale (2004) extended the Diamond-Dybvig model to include bank capital that provides additional risk sharing between risk neutral investors (equity holders) and risk averse depositors. The importance of the liquidity services provided by deposits has also been argued, more recently, for instance by Stein (2012) and De Angelo and Stulz (2015).

Our model assumes there are direct costs of default that reduce the value of the bankruptcy estate. The empirical literature shows that these costs can be substantial for both banks and non-financial firms (see James, 1991; Andrade and Kaplan, 1998; Korteweg, 2010). More recent work suggests that these estimates may understate the true costs of default (Almeida and Philippon, 2007; Acharya, Bharath and Srinivasan, 2007). We assume that these costs are true deadweight costs as distinct from the fire sale “losses” that are actually transfers of value (cf. Gale and Gottardi, 2015).

Van den Heuvel (2008) studies a quantitative model in which bank capital structure is determined by the trade-off between the liquidity services of bank debt and the costs of moral hazard that are associated with risk shifting behavior. The model does not allow for aggregate uncertainty and assumes 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 (2015) study provide a quantitative analysis of a model in which the capital structures of banks and borrowers are endogenously determined. They show that the optimal leverage in the banking sector is much higher than in the corporate sector. Banks’ assets are less risky than the borrowers’ assets for two reasons. First, the banks hold debt claims, so the first loss falls on the corporate shareholders. Second, the banks reduce the risk of their portfolio by diversifying across firms. These two factors are sufficient to produce a realistic levels of bank capital.

The paper most closely related to ours is Allen, Carletti and Marquez (2014), henceforth ACM.⁶ Like us, ACM assume 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

⁶The published version of ACM, Allen, Carletti and Marquez (2015), contains only the first part of the working paper version and deals only with banks that invest directly in projects, rather than lending to firms. The results more closely related to the present paper are found in the working paper version and, in what follows, we refer only to that version.

than equity investors, so in equilibrium the depositors receive a lower return than the equity investors. Equity is therefore an “expensive” source of funding.

The capital structure of the representative firm and bank are chosen to maximize their expected joint surplus, subject to participation constraints for the equity holders and deposit holders. This cooperative contracting approach guarantees the efficiency of the equilibrium. In our framework, by contrast, efficiency is a property of competitive equilibrium when markets are complete in the sense that all possible types of debt and equity can be traded. ACM also derive a result that bank equity has zero value in equilibrium, for the special case in which 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 default. ACM show that putting all the equity in the firms minimizes the probability of default for both banks and firms.

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.

We should also mention a large theoretical literature on the role of bank capital in preventing risk shifting or asset substitution, beginning with the seminal paper of Stiglitz and Weiss (1981) and including recent contributions such as Martinez-Miera and Repullo (2010).

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 in our model make the pricing of assets more complicated 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. See also Bisin, Gottardi and Ruta (2014) on the pricing of securities when intermediaries are present.

2 An equilibrium model of banks' capital structure

2.1 Endowments and technologies

There are two dates, indexed $t = 0, 1$, and a finite number of states of nature, $s = 1, \dots, S$. 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 $s = 1, \dots, S$.

There are two goods, a non-produced capital good and a produced consumption good. Consumption is produced subject to constant returns to scale using capital goods as the only input. There is 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 .

There is a continuum of identical consumers with unit mass. Each consumer has an initial endowment of $k_0 = 1$ units of capital at date 0. There is no initial endowment of consumption.

2.2 Firms

We assume that each active firm can invest in only one of the n technologies available. 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 is assumed to belong to a finite interval $L = [0, \ell_{\max}]$.

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. Although the number of possible types is infinite, we focus on equilibria in which the number of active types is finite.⁷

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

⁷The number of active types needed for existence is finite by Caratheodory's Theorem.

$f = (\ell, j)$ 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.

2.3 Banks

Banks lend to firms by purchasing their debt. We assume banks do not invest in firm equity.⁸ Banks raise funds by issuing deposits and equity to consumers. A bank's capital structure is determined by the level of deposits it chooses to issue. We denote the face value of deposits by $d \in D$, where D is a finite interval $[0, d_{\max}]$. The bank's portfolio is described by a vector $\mathbf{x} \in \mathbf{R}_+^F$, where $x_f \geq 0$ denotes the units of debt of type- f firms held by the bank.⁹ Since the banks' technology is subject to constant returns to scale, we can assume without loss of generality that each bank's portfolio is normalized so that $\sum_{f \in F} x_f = 1$. In other words, the bank invests in a portfolio of debt issued by firms that collectively operate one unit of capital goods. Let $X \subset \mathbf{R}_+^F$ denote the set of admissible debt portfolios.

All banks have access to all types of firms' debt and the same funding opportunities. Their portfolios and capital structures may differ, however. We refer to a bank's portfolio $\mathbf{x}$ and capital structure d as its *type*. The set of bank types is $B = X \times D$. Again, only a finite

⁸This ensures the presence of a nontrivial interdependence between the capital structure of firms and banks. Also, current regulations make this assumption realistic.

⁹In equilibrium, each bank will lend to a finite number of types of firms, since the number of active firm types is finite.

number of bank types will be active in equilibrium. 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

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

and

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

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

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

for each s . Note that the payoff vectors are completely determined by the bank's type, as in equations (3) and (4). 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, $q_b^d + q_b^e - \mathbf{q}_F^d \cdot \mathbf{x}$. In equilibrium, the maximum profit will be zero, that is, $q_b^d + q_b^e \leq \mathbf{q}_F^d \cdot \mathbf{x}$, and only banks that achieve zero profits, $q_b^d + q_b^e = \mathbf{q}_F^d \cdot \mathbf{x}$, will be active.

2.4 Consumers

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 $\beta \in (0, 1)$.¹⁰ The function

¹⁰The 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 they can be spent. This time delay reduces the value of the consumption because of discounting.

$u : \mathbf{R}_+ \rightarrow \mathbf{R}$, 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.¹¹

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) . Although consumer's have access to an infinite number of securities, the consumer's portfolio must have a finite support in equilibrium. The set of feasible portfolios is denoted by Z and defined to be the set of portfolios $\mathbf{z}$ with finite support 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

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

subject to the budget constraints,

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

2.5 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 $\boldsymbol{\kappa} = (\kappa_f)_{f \in F}$. Formally, the *allocation* is an array $(\mathbf{c}, \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\kappa})$, where $\mathbf{z}$, $\boldsymbol{\mu}$, and $\boldsymbol{\kappa}$ have finite supports. An allocation is *attainable* if

$$\sum_{f \in F} \kappa_f = 1, \tag{6}$$

$$\sum_{b \in B} \mu_b \mathbf{x} = \boldsymbol{\kappa}, \tag{7}$$

$$z_b^d = z_b^e = \mu_b, \quad \forall b \in B, \tag{8}$$

$$z_f^e = \kappa_f, \quad \forall f \in F, \tag{9}$$

¹¹ Although we maintain the assumption for simplicity, it seems quite realistic. Banks may have an advantage in monitoring firms and enforcing repayment of loans. And since loans to firms do not function as “money,” deposits are more attractive to consumers in any case.

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' endowments. 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 $\mathbf{q}$;
- (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 Constrained efficiency

In this section, we show that analogues of the Fundamental Theorems of Welfare Economics hold for the environment described above. Markets are incomplete, since banks and firms are restricted to using debt and equity, so the appropriate welfare concept is *constrained* Pareto efficiency, rather than Pareto efficiency.

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

$$\sum_{b \in B} \mu_b = 1, \quad (11a)$$

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

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) to (10). Conversely, if $(\mathbf{c}^*, \mathbf{z}^*, \boldsymbol{\mu}^*, \boldsymbol{\kappa}^*)$ is an attainable allocation, $(\mathbf{c}^*, \boldsymbol{\mu}^*)$ satisfies the constraints (11a) and (12a).¹²

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.*

Proof. See the appendix. ■

The argument of the proof is standard, and exploits the fact that markets for all the possible types of securities that can be issued by firms and banks are competitive and complete. 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.*

Proof. See the appendix. ■

Although these results look quite “standard,” they require a number of restrictive assumptions. Without a complete set of markets for contingent claims, we can only ensure equilibrium is constrained efficient. As Geanakoplos and Polemarchakis (1986) have shown, general equilibrium with incomplete markets is generically constrained inefficiency unless special conditions are satisfied. One of these conditions is the existence of a single representative consumer; another is the assumption of a single good. These assumptions are common in financial applications, but they are nonetheless 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—but the single good assumption would be harder to remove. Finally, the assumption that markets are open for all securities is crucial for the equilibrium coordination of capital structures.¹³ The resulting benchmark model

¹²To 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).

¹³Hart (1979) follows an alternative approach, in which equilibrium of a stock market economy is reached in two stages. In the first stage, firms choose production plans and have rational expectations about the value of the firm that will be realized in the second stage. This approach only requires markets to open for the shares of firms that actually form. Under sufficient regularity conditions, this appears to be equivalent to the approach adopted here.

is important for two reasons: first, it demonstrates the possibility of efficient coordination of capital structures in a decentralized economy and, second, it considerably simplifies the characterization of the equilibrium capital structures as the result of a planner's problem.

4 Banks' equilibrium capital structure

The use of equity is costly because it reduces the portion of the cash flow of banks and firms that is paid out as liquid deposits. In the case of banks, an increase in equity directly reduces the amount of deposits needed to fund the bank's portfolio. In the case of firms, an increase in equity reduces the amount firms borrow from banks and that in turn reduces the amount of funding needed by banks, either in the form of deposits or equity. The benefit of equity, of course, is that it provides a buffer against the risk of costly default. A firm's revenue is uncertain because firms invest in risky technologies and a bank's revenue is uncertain because banks lend to risky firms. In either case, the higher the size of the equity buffer, the lower the probability of default. Moreover, the role of banks as intermediaries implies that the riskiness of their portfolios depends on the size of the firms' equity buffers: the larger the firms buffer, the less likely firms are to default, and the safer is the bank's portfolio. Hence, firm equity does "double duty," in the sense that, by providing a buffer against default by the firm, it also helps to prevent default by the bank that lends to the firm.

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.

While it is clear that the cost of issuing equity is increasing in the "liquidity premium" $1 - \beta$ and the benefits are increasing in the default costs, $1 - \lambda_b$ and $1 - \lambda_f$, it is not obvious what implications this has for the optimal capital structures of banks and firms. The answer to this question depends crucially on the nature of the shocks affecting firms' technologies and the effectiveness of banks' portfolio choice in diversifying this risk, which determine how firms' productivity shocks are transmitted to banks. Uncertainty arises because the firms invest in risky technologies. This uncertainty is transmitted to banks when firms default on their loans. Firm equity does "double duty" in the sense that it protects both the firm and the bank from default. If the technology shocks are sufficiently correlated, it may be optimal for all the "loss absorption capacity" to be located in the corporate sector and none in the banking sector. Firm equity is all that is needed.

4.1 No bank equity

As we have already noted, ACM derive the surprising result that it is optimal for banks to fund themselves entirely with deposits when there is a single technology that firms can use.¹⁴

¹⁴ACM make a number of other restrictive assumptions not required in our framework: consumers are risk neutral and exogenously divided into depositors and shareholders; the single technology's productivity

We obtain an analogous result under a much weaker condition. In this section, we assume the n available technologies are *co-monotonic*.

Definition 3 *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 and we get the surprising result that, in equilibrium, banks default if *any* of their borrowers default. Each bank is so on a knife edge, with no capacity to absorb losses. It is both privately and socially optimal for banks and firms to choose capital structures that put all equity in the corporate sector. In the case of co-monotonic technologies, the double duty role of firm equity is particularly effective.

Proposition 4 *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 $b \in B$.*

It is interesting to note that this result depends only on the stochastic properties of the technology shocks. In particular, it is independent of the relative size of default costs for firms and banks. A formal proof of the proposition is found in the appendix. Here we present the main steps. Let ℓ_b denote the face value of firms' debt held by type b banks and let $\ell_b(s)$ denote the actual amount repaid to these banks by firms in state s . It is also convenient to use the notation b^* and f^* to indicate the types of banks and firms that are active in equilibrium.

Step 1: *For all active bank types $b^* = (\mathbf{x}^*, d^*)$, the value of bank equity is positive if and only if $\ell_{b^*} > d^*$.*

If $d \geq \ell_b$ it is clear that there would be nothing left over for the bank's equity holders, even if the bank's loans are repaid in full. Note that it is never optimal for an active firm to choose a face value of debt $\ell > A_{jS}$ (otherwise the firm will always be in default and incur unnecessary costs of default). As a consequence at least in the highest state, $s = S$, all firms whose debt is in the bank's portfolio are solvent and pay the face value of their debt so that, if $d^* < \ell_{b^*}$, the return to equity is $\ell_{b^*} - d^* > 0$. Limited liability ensures the payment to equity holders is non-negative in every state, so this is enough to prove that the value of equity is positive.

Step 2: *For each active firm's type $f^* = (\ell^*, j^*)$, there exists a state s_{f^*} such that firm f^* is solvent if and only if $s \geq s_{f^*}$. Similarly, for each active bank's type b^* there exists a state s_{b^*} such that bank b^* is solvent if and only if $s \geq s_{b^*}$.*

Firm f^* is solvent if and only if $A_{j^*s} \geq \ell^*$. Then the first claim follows from the fact that A_{js} is increasing in s and, as argued in Step 1, each active firm is solvent in at least one

shocks are uniformly distributed; and banks and firms choose their capital structures cooperatively.

state. Next note that the revenue of each firm (net of bankruptcy costs) is increasing in s and so the amount repaid to banks is non-decreasing in s . This, together with the fact that also for banks it is optimal to be solvent in at least one state, establishes the second claim.

Step 3: *For each active bank type b^* , the face value of deposits satisfies $d^* = \ell_{b^*}(s_{b^*})$, that is, equals the yield of the banks' portfolio in the lowest state in which the bank is solvent.*

Since bank b^* is solvent in state s_{b^*} we must have $d^* \leq \ell_{b^*}(s_{b^*})$. If $d^* < \ell_{b^*}(s_{b^*})$ the bank has the option of increasing the face value of deposits without increasing the probability of default. Since the bank is already in default in states $s < s_{b^*}$, increasing the face value of deposits will not change the amount of consumption received by deposit holders or equity holders in states $s < s_{b^*}$. In states $s \geq s_{b^*}$, on the other hand, an increase in the face value of deposits will transfer consumption to the bank's depositors from the bank's shareholders. Since one unit of consumption from equity's returns is worth β units of consumption from deposit's returns, and there is a representative consumer, so shareholders and depositors are the same individuals, this transfer will increase welfare, contradicting the constrained efficiency of equilibrium. Hence, in equilibrium, we must have $d^* = \ell_{b^*}(s_{b^*})$.

Step 4: *For all types of banks b^* that are active in equilibrium equity has no value: $d^* = \ell_{b^*}$.* Note first that it is never optimal for a bank to choose $d^* > \ell_{b^*}$ because this implies the bank defaults in each state, which contradicts what established in step 2. Suppose next that $d^* < \ell_{b^*}$, contrary to what we want to prove. Since we showed in Step 3 that $d^* = \ell_{b^*}(s_{b^*})$, this implies that at least one type f^* of firms whose debt is held by bank b^* is bankrupt in state s_{b^*} . Consider then a reduction in the face value of debt of these firms to a value equal to their revenue in state s_{b^*} . Hence the firms no longer default in state s_{b^*} and possibly in other states, so that default costs are avoided and the payment to the firms' debtholders in these states will be higher. This ensures that the bank is still solvent in state s_{b^*} and so depositors' returns are unchanged. The change will then have two effects. First, it will increase the returns both to banks' and to firms' equityholders in the states $s \in [s_{b^*}, s_{f^*}]$ where firms f^* were insolvent and are now solvent. Second, it will increase the returns to firms' equityholders and reduce, by the same amount, that of the banks' equity holders in the states $s \geq s_{f^*}$ where firms f^* were solvent. Since the equity of both firms and banks is held by the representative agent, this second effect does not affect welfare, while the first is unambiguously welfare increasing. This contradicts the constrained efficiency of equilibrium and shows so that we cannot have $d^* < \ell_{b^*}$ in equilibrium. Then the value of bank equity must be zero.

The above argument shows that default by one or more of the borrowing firms is always a *necessary* condition for the lending bank to default, because a bank will never set the face value of deposits d higher than the face value of the debt it holds. This clearly illustrates the fact that firm equity always does "double duty," serving as a buffer against both bank default and firm default.

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 ?? 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 5 *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.*

Since the bank's portfolio is endogenous, the corollary gives us no information about the conditions under which a co-monotonic portfolio will be chosen in equilibrium. It merely emphasizes that bank capital is not needed as long as the bank does not diversify its portfolio outside a set of co-monotonic technologies. In particular, if a bank does not diversify its portfolio and only lends only to firms using a single technology j , the bank will have zero equity (its portfolio is then trivially monotonic).

Proposition ?? states that banks use debt financing exclusively. It does not say anything about the capital structure of firms, however. For example, it does not claim that firms will issue equity to reduce their default risk. The firms' choice of capital structure will depend on model parameters, such as the recovery rates of banks and firms, λ_b and λ_f . The higher the default costs, other things being equal, the higher one expects the firms' equity to be. The only certainty is that banks will use no equity in equilibrium.

Proposition ?? is also silent on the variety of capital structures and technologies used by firms in equilibrium, as well as on the portfolio choice by banks. Because of the non-convexities that are an essential part of the model, we allow the full use of the convexifying effect of large numbers in order to ensure the existence of an equilibrium. Many types of firms, distinguished by their capital structures and technology choices, as well as many types of banks, distinguished by their portfolio choice, are potentially active in equilibrium. To make the analysis tractable and say more about the properties of equilibria, it will be useful to consider some special cases where the number of active types is limited. We consider a few of these cases in the next section.

To sum up, an important implication of co-monotonicity is that under this condition there is little scope for diversification in the choice of banks' portfolios. As argued above, defaults are positively correlated in the sense that an increase in the state s will reduce or leave unchanged the set of defaulting banks and firms. Thus, banks' ability to reduce their default probability by diversifying their portfolio of loans among firms of different types is quite limited. Equity is the main instrument to limit default and the efficient place to allocate equity is at the source of uncertainty, that is, in the firms at the top of the intermediation chain. Conversely, violating the *co-monotonicity assumption is a necessary condition for bank equity to have positive value in equilibrium*. For banks to issue equity, there must be benefits from diversification, that is, the possibility of reducing the probability of default by diversifying the bank's portfolio across firms using non-co-monotonic technologies. The next section identifies some environments where the benefits of diversification are present and are exploited by banks.

4.2 Positive bank equity

In this section we explore environments with both aggregate and idiosyncratic productivity shocks and identify conditions under which constrained efficiency requires a positive value of bank equity. We will assume that consumers are risk neutral, which simplifies 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.

An example with aggregate and idiosyncratic risk We consider first a simple environment in which it is possible to derive the equilibrium capital structures explicitly. There are n technologies and a finite number $S = n + 2$ of states of nature. The probability of state s is denoted by π_s and given by

$$\pi_s = \begin{cases} \frac{1-\delta-\varepsilon}{n} & \text{for } 1 \leq s \leq n \\ \delta & \text{for } s = n + 1 \\ \varepsilon & \text{for } s = n + 2 \end{cases} . \quad (13)$$

The productivity of technology j in state s is assumed to satisfy

$$A_{js} = \begin{cases} a_L & \text{if } s = j \leq n \\ a_M & \text{if } s \neq j \leq n \\ a_M & \text{if } s = n + 1 \\ a_H & \text{if } s = n + 2 \end{cases} , \quad (14)$$

where $0 < a_L < a_M < a_H$. The states are partitioned as follows:

- if a state $s \in \{1, \dots, n\}$ occurs, exactly one technology $j = s$ receives a negative shock, a_L , while the remaining technologies $j \neq s$ have normal productivity, a_M ;
- if state $s = n + 1$ occurs, all technologies have normal productivity;
- and if state $n + 2$ is realized, all technologies have a high productivity, a_H .

Since the states $1, \dots, n$ are equally likely, the technologies are ex ante identical. Also, the environment exhibits both aggregate and idiosyncratic uncertainty. We can view the realization of one of the states in the baseline set $\{1, \dots, n\}$ as a purely idiosyncratic shock, while

the realization of state $n + 1$ (respectively $n + 2$) constitutes a small (respectively large) aggregate shock relative to the baseline states.

A bank portfolio is said to be *simple* if all firms with the same technology, whose debt is held by the bank, have the same capital structure. We can show that it is optimal for banks to choose a simple portfolio when the structure of technologies satisfies (13) and (14). Since the technologies are ex ante identical, it follows that the same capital structure is optimal for all firms.

Proposition 6 *When the structure of technologies satisfies (13) and (14), 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 6 assures us that a bank will lend to firms that use only one of these capital structures. Unfortunately, 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.

The specification given by (13) and (14) incorporates a number of interesting cases. In the limit as $\delta + \varepsilon \rightarrow 1$, all technologies are identical (there is only aggregate risk) and, hence, co-monotonic. Then Proposition ?? 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 $\varepsilon \rightarrow 0$, we have the case of pure idiosyncratic risk. In each of the baseline states $\{1, \dots, n\}$, exactly one technology yields a_L and the remainder yield a_M . 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 the idiosyncratic case. These allocations are, of course, constrained efficient by Proposition ??.

Proposition 7 *Assume the technologies satisfy (13), (14) and $\delta = \varepsilon = 0$. Then in equilibrium banks' choices are as follows:*

- (i) if $\left(\frac{n-\lambda_f}{n-1}\right) a_L > a_M$, each bank lends exclusively to firms with safe debt ($\ell_j = a_L$) and¹⁵ sets its level of deposits at $d = a_L$, or
- (ii) if $\left(\frac{n-\lambda_f}{n-1}\right) a_L < a_M$, each bank lends exclusively to firms with risky debt ($\ell_j = a_M$) and chooses a fully diversified portfolio (holding $1/n$ units of the 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.

¹⁵Since the yield of the debt of firms with safe debt is always the same, whichever their technology, the composition of the bank's portfolio (with regard to the debt of these firms) is clearly irrelevant).

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.

In both of the extreme cases in our example, 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 in the efficient capital structures of both firms and banks. The following result states this more precisely.

Proposition 8 *Assume the technologies satisfy (13) and (14). Then if*

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

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 of risky firm debt*

$$\mathbf{x} = \left\{ \left(0, \frac{1}{n}, 0 \right) \right\}_{j=1}^n$$

and a face value of deposits

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

Obviously, bank equity has positive value because a bank receives a loan repayment equal to $a_M > d$ in state $s = n + 1, n + 2$. The value of firm equity is also positive 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 8 is that the size of the negative idiosyncratic shock, $a_M - a_L$, is sufficiently large, so that, when $\delta = \varepsilon = 0$, firms choose $\ell = a_M$ and banks choose a fully diversified portfolio in equilibrium. This corresponds to Case (ii) of Proposition 7. 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 8, 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. On the other hand, the introduction of an equity buffer to prevent the default of a diversified bank would be much cheaper. 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 table below summarizes the equilibrium values of bank and firm equity in the different cases characterized by Propositions 7 and 8:¹⁶

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

Note that whether firm equity has positive or zero value in the co-monotonic case $\varepsilon + \delta = 1$ 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 (13) and (14) also allow us to explore the connection between portfolio diversification and the presence of positive bank equity. We noted at the end of Section ?? 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 7, where bank equity has zero value and the bank's portfolio is risk free. But there are other situations 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 9 *Assume that firms' technologies satisfy (13), (14) 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 ; firms using technology j have face value of debt $\ell_j = a_M$, and bank deposits are $d = a_M$. Both the firms using technology j and the banks lending to them default in state $s = j$.*

The proof follows an argument parallel to the proof of Proposition 8. 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 9 holds. To avoid default, even with diversification, an equity buffer is needed, as we saw in Proposition 8, 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 probability of default is minimized by choosing a simple portfolio and the benefit of this choice increases as $\delta \rightarrow 1$. Thus, even if diversification is possible, it is not always optimal. This principle is also illustrated by ACM, who analyse an environment with two i.i.d. technologies and

¹⁶We maintain here the condition $\left(\frac{n-\lambda_L}{n-1}\right) a_L < a_M$.

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 the bank equity to have zero value, even though there appears to be scope for diversification. 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, for bank equity to have positive value.

A “single factor” model 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. Single factor models have also been incorporated in equilibrium models, such as Martinez-Miera and Repullo (2010), in order to study the effect of capital requirements on bank failure. In single factor models, the value of the assets $A_j(T)$ of a borrowing firm j at time T are assumed to be the product of two random variables,

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

where z is a standard normal aggregate shock, ε_j is a standard normal idiosyncratic shock, ρ is the correlation parameter, and the random variables z and $\{\varepsilon_j\}$ are i.i.d.. 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.

We can obtain something very similar in our model if we assume that there is a large number of technologies $j \in [0, 1]$ with productivities per unit of capital given by $A_j = \theta_j a$, where θ_j and a are random variables with c.d.f.s F and G , respectively. Both F and G are C^2 and satisfy

$$F(\theta_0) = 0, F(\theta_1) = 1, \text{ and } F'(\theta_j) > 0 \text{ for } \theta_0 < \theta_j < \theta_1, F'(\theta_0) = 0 = F'(\theta_1) \quad (15)$$

and

$$G(a_0) = 0, G(a_1) = 1, \text{ and } G'(a) > 0 \text{ for } a_0 < a < a_1. \quad (16)$$

The random variables θ_j and a represent idiosyncratic and aggregate shocks, respectively. The random variables $\{\theta_j\}$ are assumed to be i.i.d and independent of the aggregate shock a . We also normalize the variables θ_j so that $E[\theta_j] = 1$. The law of large numbers holds, so the cross sectional distribution of the idiosyncratic shock is given by F with probability one. It is clear that the available technologies are not co-monotonic, though the productivities are positively correlated because of the common shock a .

Our main focus here is on the capital structure decisions by banks. To this end, and also in order to mimic the properties of single factor models, which take capital structures and the loan portfolio as exogenous and (sometimes) symmetric, we assume here that all firms choose the same capital structure, regardless of their technology. We can argue this is natural since the technologies are ex ante identical.¹⁷ We also assume that every bank

¹⁷Technically, the productivities of the different technologies are *exchangeable* random variables, that is, the joint distribution function is independent of any permutation of the random variables. Hence we can argue the environment is similar to the one considered at the beginning of this section, where we showed that all firms choose the same capital structure and all banks choose simple portfolios.

chooses a fully diversified and symmetric portfolio containing an equal amount of the debt issued by firms choosing each technology j .

When the face value of firms' debt is given by ℓ , the return on a fully diversified portfolio of firms' debt depends only on the aggregate state a and is denoted by $\ell(a)$. We need to consider three cases, depending on the size of the aggregate productivity shock. In the first case, the aggregate shock a is sufficiently small that all firms default. This happens if and only if $\theta a < \ell$ for all $\theta_0 \leq \theta \leq \theta_1$, which is equivalent to $\theta_1 a \leq \ell$ or $\theta_1 \leq \ell/a$. In that case, every firm pays the bank $\lambda_f \theta a$ and the total repayment is

$$\ell(a) = \int_{\theta_0}^{\theta_1} \lambda_f \theta a dF = \lambda_f a,$$

since $\mathbf{E}[\theta] = 1$. In the second case, firms with the lowest idiosyncratic shocks default and firms with the highest idiosyncratic shocks repay in full. This case occurs if $\theta a = \ell$ for some $\theta_0 < \theta < \theta_1$, or $\theta_0 < \ell/a < \theta_1$. Firms with a shock $\theta < \ell/a$ default and repay $\lambda \theta a$, firms with a shock $\theta \geq \ell/a$ repay ℓ , and the total repayment to the bank is

$$\ell(a) = \int_{\theta_0}^{\ell/a} \lambda_f \theta a dF + \ell \left(1 - F\left(\frac{\ell}{a}\right) \right).$$

In the third case, the aggregate shock is high enough that all firms are solvent and can repay the face value of the debt in full. This happens if $\theta a \geq \ell$ for all $\theta_0 \leq \theta \leq \theta_1$, which is equivalent to $\theta_1 \geq \ell/a$. Each firm repays ℓ and the total repayment is $\ell(a) = \ell$. Thus, the return on the bank's portfolio is

$$\ell(a) = \begin{cases} \lambda_f a & \text{if } \theta_1 \leq \ell/a \\ \lambda_f \int_{\theta_0}^{\ell/a} \theta a dF + \ell (1 - F(\ell/a)) & \text{if } \theta_0 < \ell/a < \theta_1, \\ \ell & \text{if } \ell/a \leq \theta_0 \end{cases}.$$

We see from the expression of $\ell(a)$ that in this single factor model the return on the bank's portfolio depends only on the aggregate state a , but idiosyncratic risk plays a role in determining the size of the total repayment to the bank. This is clearly seen in the second case considered above, where a fraction $F(\ell/a)$ default and a fraction $1 - F(\ell/a)$ repay in full. Because of the default cost, $1 - \lambda_f$, the larger the number of defaults, the lower the repayment to the bank, other things being equal. Although diversification eliminates idiosyncratic risk, the presence of idiosyncratic risk matters when there is default. Note that in the first two cases, the repayment function is increasing in a . In the third case, the repayment function is independent of a . Also under the stated assumptions on F , both $\ell(a)$ and $\ell'(a)$ are continuous.¹⁸

We proceed now to characterize the conditions under which bank equity is positive in equilibrium. Let a_b denote the smallest value of $a \geq a_0$ such that $\ell(a) \geq d$, that is, a bank issuing debt d is solvent. Note that it will never be optimal for the bank to choose $d > \ell(a)$

¹⁸See the proof of Proposition 11 in the appendix for a formal demonstration.

for all values of a , as that would mean that banks always default, which is suboptimal. Hence a_b is well defined and $a_b \leq a_1$. The bank will also be solvent for all $a > a_b$, because $\ell(a)$ is monotonically non-decreasing, as we noted above. Similarly, let a_f denote the smallest value of $a \geq a_0$ such that no firm defaults, that is, $\theta a_f \geq \ell$ for all $\theta_0 \leq \theta \leq \theta_1$, or $a_f \geq \ell/\theta_0$. If $\ell/\theta_0 \leq a_1$ then it is clear that $a_f = \ell/\theta_0$. Otherwise, a_f is not well defined and, in that case, we set $a_f = \infty$ by convention. We then show the following:

Lemma 10 *The equilibrium value of bank equity is positive if and only if $a_b < \min\{a_f, a_1\}$.*

That is, bank equity has a positive value if the bank is solvent in more than one aggregate state and defaults in fewer states than any firm. Using this result, we can show that it is never optimal to have the value of bank equity equal to zero.

Proposition 11 *If the bank technologies satisfy (15) and (16), all firms choose the same capital structure, independently of the technology j , and banks lend equally to firms using each technology, then the value of bank equity is strictly positive.*

The proof shows that if bank equity were zero, a reduction in the face value of banks' debt, by reducing banks' default costs always increases their market value. Hence in this single factor environment, where diversification allows banks to reduce, but not completely eliminate the risk in their portfolio, bank equity has a comparative advantage over firm equity.

5 Conclusion

We have presented a classical model of competitive equilibrium in which banks act as intermediaries between productive firms and consumers. 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 bank, is determinate in equilibrium. 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, so that risk and equity capital are optimally allocated between the two sectors.

We have studied an environment in which the laissez-faire equilibrium is constrained efficient and there is no obvious need for capital adequacy regulation. Nonetheless, this may be considered a first step toward a theory of bank capital regulation. In the first place, the characterization of the constrained efficient capital structures gives us some idea of what

capital adequacy regulation should lead to. In the second place, contrasting the conditions under which *laissez faire* is constrained efficient with the real world gives us a clue why market failures occur. In particular, our assumption that perfect markets exist for all possible types of debt and equity is quite demanding.

We alluded briefly in the introduction to some of the justifications for 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 moral hazard in the form of risk shifting and asset substitution. The possibility of government bailouts in the event of default may also 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 many 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 the fact that policymakers have focused on the role of bank capital as a buffer, rather than an incentive mechanism. Bank capital provides a cushion that shields depositors and bond holders from losses but, more importantly from the point of view of politicians, it also makes politically unpopular bailouts unnecessary. Whether innovations such as Total Loss Absorption Capacity and bail-inable debt will actually put an end to bailouts is not clear, but capital regulation motivated by the desire to avoid bailouts will 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] Admati, A. and M. Hellwig (2013): *The bankers' new clothes: What's wrong with banking and what to do about it*. Princeton University Press: Princeton and Oxford.
- [3] Allen, F., E. Carletti, and R. Marquez (2014): Deposits and Bank Capital Structure, unpublished working paper.
- [4] Allen, F., E. Carletti, and R. Marquez (2015): Deposits and Bank Capital Structure. *Journal of Financial Economics*, **118**, pp. 601-19

- [5] Allen, F. and D. Gale (1988): Optimal Security Design. *Review of Financial Studies*, **1**, pp. 229-263
- [6] Allen, F. and D. Gale (1991): Arbitrage, Short Sales, and Financial Innovation. *Econometrica*, **59**, pp. 1041-1068
- [7] Almeida, H. and T. Philippon (2007): The risk-adjusted cost of financial distress. *Journal of Finance*, **62**, pp. 2557–2586
- [8] 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
- [9] 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.
- [10] Bisin, A., P. Gottardi, and G. Ruta (2014): Equilibrium corporate finance and intermediation, NBER Working Paper No. 20345.
- [11] 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.
- [12] 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.
- [13] Cihak, M., A. Demirguc-Kunt, E. Feyen, and R. Levine (2012): Benchmarking financial systems around the world. Unpublished policy research working paper 6175. World Bank, Washington, DC.
- [14] 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.
- [15] 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
- [16] 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.
- [17] Diamond, D. (1984): Financial intermediation and delegated monitoring. *Review of Economic Studies*, **60**, pp. 393–414.
- [18] Diamond, D., and P. Dybvig (1983): Bank runs, deposit insurance and liquidity. *Journal of Political Economy*, **91**, pp. 401–419.

- [19] 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.
- [20] 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.
- [21] 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
- [22] 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.
- [23] Gale, D. and P. Gottardi (2015): Capital Structure, Investment and Fire Sales. *Review of Financial Studies*, **28**, pp. 2502-33.
- [24] Gornall, W. and I.A. Strebulaev (2015): Financing as a Supply Chain: The Capital Structure of Banks and Borrowers, Stanford University Working Paper.
- [25] Gropp, R. and F. Heider (2010): The determinants of capital structure: some evidence from banks. *Review of Finance*, **14**, pp. 587–622.
- [26] Grossman, S.J. and O.D. Hart (1979): A theory of competitive equilibrium in stock market economies, *Econometrica*, **47**, pp. 293-329.
- [27] Hackbarth, D., and D. C. Mauer (2012): Optimal priority structure, capital structure, and investment. *Review of Financial Studies*, **25**, pp. 747–96.
- [28] James, C. (1991): The losses realized in bank failures. *Journal of Finance*, **46**, pp. 1223–1242
- [29] Kim, E. H. (1982): Miller’s equilibrium, shareholder leverage clienteles, and optimal capital structure. *Journal of Finance*, **37**, pp. 301–19.
- [30] Korteweg, A. (2010): The net benefits to leverage. *Journal of Finance*, **65**, pp. 2137–2170
- [31] 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.
- [32] Makowski, L. (1983): Competitive stock markets. *Review of Economic Studies*, **50**, pp. 305-330.
- [33] Martinez-Miera, D. and R. Repullo (2010): Does competition reduce the risk of bank failure? *Review of Financial Studies*, **23**, pp. 3638-64.

- [34] Mehran, H. and A. Thakor (2011): Bank capital and value in the cross section. *Review of Financial Studies*, **24**, pp. 1019–1067.
- [35] Modigliani, F. and M. Miller (1958): The cost of capital, corporation finance, and the theory of investment *American Economic Review*, **48**, pp. 261–297.
- [36] Radner, R. (1974): A note on 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.
- [37] Stein, J.C. (2012): Monetary policy as financial stability regulation, *Quarterly Journal of Economics*, *127*, pp. 57–95.
- [38] Stiglitz, J.E. and A. Weiss (1981): Credit rationing in markets with imperfect information, *American Economic Review*, **71**, pp. 393–410.
- [39] Titman, S. (1984): The effect of capital structure on a firm’s liquidation decision. *Journal of Financial Economics*, **13**, pp. 137–51.
- [40] Titman, S., and R. Wessels (1988): The determinants of capital structure choice. *Journal of Finance*, **43**, pp.1–19.
- [41] Van den Heuvel, S. (2008): The welfare cost of bank capital requirements. *Journal of Monetary Economics*, **55**, pp. 298–320
- [42] Vasicek, O. (2002): Loan Portfolio Value. *Risk*, **15**, pp. 160–62.

Appendix

Proof of Proposition ??

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

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

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 $\sum_{b \in B} \mu_b x_f = \kappa_f$, 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 ??

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 $\mathbf{x}$,

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

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 ??

In state s , the holders of equity in bank $b = (\mathbf{x}, d)$ receive [WHY RESTRICT THE SUM ONLY TO FIRMS WHO ARE ACTIVE IN EQUILIBRIUM?]

$$\max \left\{ \sum_{f^*=(\ell^*, j^*)} \chi_{f^*}(s) x_{f^*} \min \{\ell^*, A_{f^*s}\} - d, 0 \right\}$$

units of consumption and the holders of deposits receive

$$\chi_b(s) \min \left\{ d, \sum_{f^*=(\ell^*, j^*)} \chi_{f^*}(s) x_{f^*} \min \{\ell^*, A_{f^*s}\} \right\}$$

units of consumption, where

$$\chi_{f^*}(s) = \begin{cases} 1 & \text{if } \ell^* \leq A_{j^*s} \\ \lambda_{f^*} & \text{if } \ell^* > A_{j^*s} \end{cases}$$

and

$$\chi_b(s) = \begin{cases} 1 & \text{if } d \leq \sum_{f^*=(\ell^*, j^*)} \chi_{f^*}(s) x_{f^*} \min \{\ell^*, A_{j^*s}\} \\ \lambda_{f^*} & \text{if } d > \sum_{f^*=(\ell^*, j^*)} \chi_{f^*}(s) x_{f^*} \min \{\ell^*, A_{j^*s}\} \end{cases}.$$

Let

$$\ell_b = \sum_{f^*=(\ell^*, j^*)} x_{f^*} \ell^*$$

denote the face value of the debt owed to bank b and let

$$\ell_b(s) = \sum_{f^*=(\ell^*, j^*)} \chi_{f^*}(s) x_{f^*} \min \{\ell^*, A_{j^*s}\}$$

denote the actual amount repaid to bank b in state s .

Step 1: *The value of equity of all active banks 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 $\ell_b(s) > d$ in at least one state s . Since default is costly, all active firms choose a debt level $\ell^* \leq A_{fS}$ and all active banks a level of deposits $d^* \leq \ell_{b^*}$. We claim that the value of equity of bank b^* is positive if and only if $\ell_{b^*} > d^*$. 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^* repays the face value of its debt, the bank receives ℓ_{b^*} and the equity holders receive $\ell_{b^*}(S) - d^* = \ell_{b^*} - d^* > 0$.

Step 2: *For each f^* , there exists a state s_{f^*} such that firm f^* is solvent if and only if $s \geq s_{f^*}$. Similarly, for each b^* 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 $\ell^* > A_{j^*s}$. 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_{j^*s} 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 $\ell_{b^*}(s)$ is non-decreasing in s because A_{j^*s} is increasing in s for every j and $\chi_{f^*}(s)$ is non-decreasing in s , for every f^* . Let

$$d_{b^*}(s) = \chi_{b^*}(s) \min \{d^*, \ell_{b^*}(s)\}$$

denote the payment to deposit holders in state s . Then it is clear that $d_{b^*}(s)$ is non-decreasing in s because $\ell_{b^*}(s)$ and $\chi_{b^*}(s)$ are non-decreasing. 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 $d^* = \ell_{b^*}(s_{b^*})$.*

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 $d^* < \ell_{b^*}(1)$. Then increasing d to $d^* + \varepsilon < \ell_{b^*}(1)$, say, will increase the payout to depositors by ε in every state and reduce the payout to equity holders by the same amount. This increases effective consumption by $(1 - \beta)\varepsilon$ in each state, contradicting the constrained efficiency of the equilibrium.

The second possibility is that $s_{b^*} > 1$ and $d^* < \ell_{b^*}(s_{b^*})$. Suppose the bank increases deposits by $\varepsilon > 0$ to $d^* + \varepsilon < \ell_{b^*}(s_{b^*})$. This will not have any effect in states $s < s_{b^*}$ because the bank is in default, depositors are receiving $\ell_{b^*}(s) < d^*$ and 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 $(1 - \beta)\varepsilon$. This again contradicts constrained efficiency.

Step 4: *Bank b 's equity has no value: $d^* = \ell_{b^*}$.*

Suppose to the contrary that $d^* = \ell_{b^*}(s_{b^*}) < \ell_{b^*}$. Then there must exist at least one firm that is bankrupt in state s_{b^*} . Otherwise, the firms' repayment would be $\ell_{b^*}(s_{b^*}) = \ell_{b^*}$.

Suppose that firm $f^* = (\ell^*, j^*)$ is bankrupt in state s_{b^*} and consider the effect of reducing the borrowing of f^* by an amount $\varepsilon > 0$ such that $\ell^* - \varepsilon = A_{j^*s_{b^*}}$. This change has no impact on the viability of the bank in state s_{b^*} because firm f^* is now paying $A_{j^*s_{b^*}}$ in state s_{b^*} , instead of $\lambda_{f^*}A_{j^*s_{b^*}}$. 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's debt (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 equity of the bank. In states $s \geq s_{b^*}$, the payoff of bank's equity will increase because of the increase in the firm's repayment, that is the change in the payoff of bank's equity is equal to

$$\Delta_{b^*,s} = \begin{cases} A_{j^*s_{b^*}} - \lambda_{f^*}A_{j^*s} & \text{if } s_{b^*} \leq s < s_{f^*} \\ A_{j^*s_{b^*}} - \ell^* & \text{if } s_{f^*} \leq s \end{cases},$$

because the firm will pay $A_{j^*s_{b^*}}$ in each state and was previously paying $\lambda_{f^*}A_{j^*s}$ in states $s_{b^*} \leq s < s_{f^*}$ and ℓ^* in states $s_{f^*} \leq s$.

The return to the equity of firm f^* is increased by $A_{j^*s} - A_{j^*s_{b^*}}$ for all $s \geq s_{b^*}$; the change in the firm's equity is so

$$\Delta_{f^*,s} = \begin{cases} A_{j^*s} - A_{j^*s_{b^*}} & \text{if } s_{b^*} \leq s < s_{f^*} \\ \ell^* - A_{j^*s_{b^*}} & \text{if } s_{f^*} \leq s \end{cases}.$$

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 $\Delta_{f^*,s}$, which is

$$\begin{aligned} \Delta_{b^*,s} + \Delta_{f^*,s} &= \begin{cases} A_{j^*s_{b^*}} - \lambda_{f^*}A_{j^*s} + A_{j^*s} - A_{j^*s_{b^*}} & \text{if } s_{b^*} \leq s < s_{f^*} \\ A_{j^*s_{b^*}} - \ell^* + \ell^* - A_{j^*s_{b^*}} & \text{if } s_{f^*} \leq s \end{cases} \\ &= \begin{cases} A_{j^*s} - \lambda_{f^*}A_{j^*s} & \text{if } s_{b^*} \leq s < s_{f^*} \\ 0 & \text{if } s_{f^*} \leq s \end{cases}. \end{aligned}$$

Thus, the net gain for equity holders is the saving in default costs $(1 - \lambda_{f^*})A_{j^*s}$ 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.

Proof of Proposition 6.

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 $\sum_j x_j^L = 1$. 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^L 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 $\sum_{j=1}^n a_j^H + a_j^M = 1$ and that the amounts invested in firms with technology j and ℓ_j equal to a_H and a_M are $a_j^H(1 - \gamma)$ and $a_j^M(1 - \gamma)$, 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 number of firms of type j with $\ell_j = a_H$, x_j^M is the number 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

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

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

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

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

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

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

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

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

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

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 7.

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

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

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

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

Multiplying by $\frac{n}{n-1}$ yields

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

and collecting like terms gives us

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

This can be rewritten as

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

which is equivalent to

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

or

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

Proof of Proposition 8

The argument in the proof of Proposition 6 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

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

- 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

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

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

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

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

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

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

Proof of Lemma 10

We first note that it is impossible to have $a_b > a_f$, because this implies that $d > \ell(a_f) = \ell$, that is, banks always default, which is suboptimal. This, together with the property $a_b \leq a_1$, established in the text, shows that we always have $a_b \leq \min\{a_f, a_1\}$.

Proof. Sufficiency. Suppose that $a_b < \min\{a_f, a_1\}$. As shown in the proof of Proposition 11 $\ell(a)$ is strictly increasing on the interval $[a_b, a_f]$, hence $\ell(a) > d$ on the interval $(a_b, a_1]$, so the value of bank equity is positive because

$$\int_{a_0}^{a_1} \max\{\ell(a) - d, 0\} dG = \int_{a_b}^{a_1} (\ell(a) - d) dG > 0.$$

Necessity. If $a_b = a_f \leq a_1$, then $d = \ell(a_b)$, $\ell(a_f) = \ell$ and $d = \ell$, so obviously bank equity has zero value.

On the other hand, $a_b = a_1 \leq a_f$ implies that $d = \ell(a_1)$, and the fact that $\ell(a)$ is monotonically non-decreasing implies that

$$\int_{a_0}^{a_1} \max\{\ell(a) - d, 0\} dG \leq \max\{\ell(a_1) - d, 0\} = 0.$$

This completes the proof. ■

Proof of Proposition 11

By direct calculation, we show that $\ell(a)$ is strictly increasing when $\theta_0 < \frac{\ell}{a} < \theta_1$:

$$\begin{aligned} \ell'(a) &= \lambda_f \int_{\theta_0}^{\ell/a} \theta dF + \lambda_f \ell F' \left(\frac{\ell}{a} \right) \frac{-\ell}{a^2} - \ell F' \left(\frac{\ell}{a} \right) \frac{-\ell}{a^2} \\ &= \lambda_f \int_{\theta_0}^{\ell/a} \theta dF + (1 - \lambda_f) F' \left(\frac{\ell}{a} \right) \left(\frac{\ell}{a} \right)^2 > 0. \end{aligned}$$

Moreover,

$$\begin{aligned} \lim_{a \rightarrow \ell/\theta_1} \ell'(a) &= \lambda_f \int_{\theta_0}^{\theta_1} \theta dF + (1 - \lambda_f) F'(\theta_1) (\theta_1)^2 = \lambda_f, \\ \lim_{a \rightarrow \ell/\theta_0} \ell'(a) &= (1 - \lambda_f) F'(\theta_0) (\theta_0)^2 = 0 \end{aligned}$$

where the final equalities follow from our assumption that $\int_{\theta_0}^{\theta_1} \theta dF = 1$ and that $F'(\theta_0) = F'(\theta_1) = 0$. This shows that $\ell'(a)$ is continuous and well defined at $\ell/a = \theta_0$ and $\ell/a = \theta_1$.

Suppose that $a_b = a_f \leq a_1$. Note that, by the definition of a_f we have $\theta_0 a_f = \ell = \ell(a_f)$, while from the continuity of $\ell(a)$ we get $\ell(a_b) = d$. We show in what follows that a small decrease in d to d' (holding ℓ constant), corresponding to a decrease in a_b to a' , always increases bank profits. At a' we have

$$d' = \ell(a') = \lambda_f \int_{\theta_0}^{\ell/a'} \theta a' dF + \ell \left(1 - F \left(\frac{\ell}{a'} \right) \right),$$

and, by the property established above,

$$\lim_{a' \rightarrow a_b = \ell/\theta_0} \ell'(a') = 0.$$

Consider then the market value of a bank with debt level $d' = \ell(a')$:

$$\lambda_b \int_{a_0}^{a'} \ell(a) dG + \ell(a') (1 - G(a')) + \beta \left[\int_{a'}^{a_f} (\ell(a) - \ell(a')) dG + (\ell - \ell(a')) (1 - G(a_f)) \right].$$

Differentiating with respect to a' and evaluating the derivative at $a' = a_f$ yields

$$\begin{aligned} & \lambda_b \ell(a') G'(a') + \ell'(a') (1 - G(a')) - \ell(a') G'(a') - \beta (\ell(a') - \ell(a')) G'(a') \\ & - \int_{a'}^{a_f} \ell'(a') dG - \beta \ell'(a') (1 - G(a_f)) - \beta (\ell - \ell(a')) G'(a_f) \\ = & \lambda_b \ell(a_f) G'(a_f) + \ell'(a_f) (1 - G(a_f)) - \ell(a_f) G'(a_f) - \beta \ell'(a_f) (1 - G(a_f)) - \beta (\ell - \ell(a_f)) G'(a_f) \\ = & (\lambda_b \ell(a_f) - \beta \ell - (1 - \beta) \ell(a_f)) G'(a_f) + (1 - \beta) \ell'(a_f) (1 - G(a_f)) \\ = & (\lambda_b - 1) \ell G'(a_f) < 0, \end{aligned}$$

because $\ell'(a_f) = 0$ and $d = \ell = \ell(a_f)$. Thus, a small decrease in d and a_b holding ℓ constant must increase the bank's market value, contradicting the equilibrium conditions.

It remains to consider the case $a_f > a_1$ so that $\ell/a_1 > \theta_0$ and $a_b = a_1$. In that case, the bank's debt level is any $d \geq \ell(a_1)$. The bank's market value is

$$\lambda_b \int_{a_0}^{a_1} \ell(a) dG$$

If we consider again a marginal decrease of a_b to a' (equivalently, of d below $\ell(a_1)$), its market value changes as follows:

$$\lambda_b \int_{a_0}^{a'} \ell(a) dG + \ell(a') (1 - G(a')) + \beta \int_{a'}^{a_1} (\ell(a) - \ell(a')) dG.$$

Differentiating again this expression with respect to a' and evaluating then the derivative at $a' = a_1$ yields:

$$\begin{aligned} & \lambda_b \ell(a') G'(a') + \ell'(a') (1 - G(a')) - \ell(a') G'(a') - \\ & \beta (\ell(a') - \ell(a')) G'(a') - \int_{a'}^{a_1} \ell'(a') dG \\ = & \lambda_b \ell(a_1) G'(a_1) + \ell'(a_1) (1 - G(a_1)) - \ell(a_1) G'(a_1) \\ = & (\lambda_b \ell(a_1) - \ell(a_1)) G'(a_1) + \ell'(a_1) (1 - G(a_1)) \\ = & (\lambda_b - 1) \ell(a_1) G'(a_1) < 0, \end{aligned}$$

In this case too, the banks' market value can be increased by lowering d below $\ell(a_1)$, that is, by having a positive value of equity.