Default, Bailouts and the Vertical Structure of Financial Intermediaries *

Tatiana Damjanovic[†] Durham University

Vladislav Damjanovic[‡] Durham University

Charles Nolan[§] University of Glasgow

August 2017

Abstract

Vertical integration of deposit-taking institutions with those investing in risky equity is studied. Leveraged financial intermediaries optimise default given government bailouts. The planner's problem identifies wedges of inefficiency in decentralized equilibrium. Societal benefits of integration flow from elimination of a credit spread reflecting a monopolistic mark-up and a risk premium, which increases labour demand and output. Costs include larger, and more frequent, bailouts. The optimal structure of intermediaries depends on shocks hitting the economy, competitiveness of the financial sector and the efficiency of government intervention. Separated institutions are preferred when profit margins are small and financial shocks are volatile and uncorrelated across institutions. Deposit insurance boosts output regardless of intermediaries' structure.

JEL Classification: E13, E44; G11; G24; G28.
Keywords: Financial intermediation in DSGE models, Vertical structure of financial intermediary, separation of retail and investment banks, bailouts, deposit insurance.

*Acknowledgements: Earlier versions of this work have benefitted from comments from seminar participants at the Bank of England, Sveriges Riksbank, the Universities of Birmingham, Cardiff, Durham, Heriot-Watt, Lancaster, Leicester, Loughborough, Manchester, NYU, Sheffield, St Andrews, CEF 2011 meetings in San Francisco, 2011 European Meeting of the Econometric Society, 2012 Meeting of the Royal Economic Society, Policy and Growth Conference, Durham University 2013, ESRC 2013 Conference Financial Modelling Post-2008: Where Next?, 2015 Macro Workshop at the University of Bath, and Boyan Jovanovic, Douglas Gale, Mark Gertler, Martin Ellison, Harris Dellas, Roberto Billi, Jesper Linde, Tony Yates, Christoph Thoenissen and Dudley Cooke. The usual disclaimer applies.

[†]Tel +44 (0) 191 33 45198. E-mail: Tatiana.Damjanovic@durham.ac.uk.

[‡]Tel +44 (0) 191 33 45140. E-mail: Vladislav.Damjanovic@durham.ac.uk.

[§]Tel +44 (0) 141 330 8693. E-mail: Charles.Nolan@glasgow.ac.uk

1 Introduction

The recent financial crisis dramatically reaffirmed that financial instability can induce macroeconomic instability. Similar experience in the past led some to recommend partitioning financial intermediaries into safer and riskier entities and adjusting regulatory practice appropriately. Some proposals were quite radical, but policymakers over time appeared largely to step back from wide-ranging structural reforms. Following the recent crisis, restructuring policies are again being introduced or considered[1].

This paper considers a policy which insulates retail deposits from investment in risky equities, similar to aspects of the Glass Steagall Act and the wider response (e.g., deposit insurance) to the Great Depression[2]. There are few macroeconomic models in the literature appropriate for assessing such reforms. This paper is an attempt to begin filling that gap. In particular we will consider vertical interaction between risky investment and commercial banking, which raises deposits and invests in loans. We will call 'investment banking' the downstream part of financial intermediation which directly finances risky entrepreneurs through the purchases of their equities. The investment banks fund their equity stake by raising loans from the retail banking sector. Retail banks, the safer part of financial intermediation, are funded by private agents' deposits. Initially we set out the problems facing the retail banks and the investment banks separately; we then 'merge' these institutions to model the implications of universal banking and compare this to the planning solution.

The model has three distinctive features. First, investment banks have projects with uncertain returns; in effect they take leveraged equity stakes in intermediate goods producers. They choose the profit maximizing level of borrowing before demand conditions are known and hence choose the likelihood of their defaulting. Second, we assume that both retail and investment banks enjoy a form of limited liability in the sense that, if they make a loss, they are allowed to continue trading next period without carrying over that loss. An alternative description of the environment is that banks go bust and are replaced next period by new banks such that market structure is identical period-to-period. What

[1]In Europe there are the Liikanen proposals. In the UK there are the Vickers proposals and in the US there is the Volcker Rule. The similarities and differences between these have been the subject of an important debate which lies beyond the scope of the present paper. Earlier proposals by Milton Friedman and others are even more radical.

[2]Other dimension of the Glass Steagall Act are discussed in Boot and Thakor (1997), who consider a merger between equity underwriting and loan provision leaving deposit holding issues to one side. Those services are substitutable and so their focus is more on a horizontal type of merger. The result is intuitive: horizontal integration reduces the size of financial services.

is key, is that banks' optimization strategies are affected by limited liability: it encourages *more* risk-taking by banks, which, other things constant, boosts the size of investment and narrows the credit spread. Finally, there is a rich menu of shocks. Investment banks are subject to idiosyncratic shocks. They also face a shock that is common to all investment banks. Hence, depending on whether common or idiosyncratic shocks are dominant, the economy may be well-insured against, or vulnerable to, financial shocks. In addition, there is a common macroeconomic shock similar to a TFP or quality of financial capital shock, as in Gertler and Kiyotaki (2010).

1.1 The vertical structure of financial intermediaries

It is important to emphasize that the welfare assessment of vertical separation of risky investment from deposit holding needs to be conducted in a general equilibrium environment. That is because one needs to analyse the costs and benefits of increasing risk. Higher risk is concomitant with greater credit availability and larger overall output. However, elevated risk makes the banking sector more fragile and may impose a larger burden on the public finances when deposit insurance is bankrolled by the taxpayer. Our model is sufficiently rich to begin such an analysis. Spengler (1950) showed that vertical integration reduces inefficiency as it eliminates double marginalization. The relevant margin in the financial sector is the credit spread: in addition to a monopolistic mark up, which will vary in the degree of competition, it also includes a risk premium. That premium depends on wider economy risks. Hence the benefit from vertical integration is different from other industries because of the possibility of default. As we discuss in more detail below, when common shocks to investment banks dominate, universal banking may be the preferred structure (see Benston, 1994). However, as we explain presently, despite the boost in output, universal banking turns out to result in larger and more frequent bailouts. These bailouts entail an excess burden: One may think of this as an agency cost (as in Carlstrom and Fuerst, 1997) or the cost associated with government intervention (such as a deadweight loss from taxation). This trade-off between the benefits from larger investment and the costs of banking fragility is one of our main findings. We also provide a detailed analysis of the planning problem further to clarify the wedges of inefficiency in the model.

The universal bank, whilst making more loans and receiving larger profits than a separate investment bank, is also more fragile. That is for the following reason. Although they charge the same mark up, the marginal cost of funds is larger for separated investment banks. As a consequence separated investment banks encounter lower demand and lower profits. However, the *sum* of profits of separated commercial and investment banks is larger than the profits of the universal bank. Other things constant, separated institutions have jointly larger quantities of own funds.

An additional focus of our research is the assessment of the benefits from deposit protection which we assume, realistically, is provided by government. The seminal paper of Diamond and Dybvig (1983) provided a rationale for deposit insurance as a way to prevent bank runs. We show that deposit protection may have another important function; it induces an increase in the labour supply and boosts output. However, the net welfare effect of deposit protection depends on how efficient the government is in raising funds. When the cost is small, we show that a bailout may indeed be welfare enhancing.

Our analysis of vertical integration suggests a key trade-off exists between a higher cost of borrowing (i.e., double marginalization leads to a larger interest rate on loans) and relatively low default when banks are separated, against more competitive pricing and larger government bailouts under universal banking[3]. A higher profit margin in the financial sector reflects higher costs of investment and shrinks the economy. On the other hand, that higher margin reduces the probability and size of government intervention when an adverse shock hits the economy. Hence, there is a trade off between the benefits of larger investment and the costs of bailouts. The evaluation of this trade off is only possible within a general equilibrium model. To the best of our knowledge, the present paper is the first to provide welfare analysis of the benefits of the separation between deposit holding and direct investment in equity.

The welfare assessment of the structure of banks and their effect on the macroeconomy is carried out analytically, via formulation of the planner's problem. It implies that optimal policy needs to be predicated on: (i) the kinds of shock hitting the economy; (ii) the degree of competitiveness of the investment and retail banking sectors and (iii) the efficiency of government intervention. It is not clear that current policy proposals are securely grounded in the issues our analysis suggests are central to optimal policy evaluation.

1.2 Related literature

By introducing default, this paper seeks to extend a burgeoning literature that incorporates forms of financial intermediation into workhorse macroeconomic models; Brunnermeier,

[3] Our results are in line with Boyd, Chang and Smith (1998) who show that universal banking structure requires a larger FDIC.

Eisenbach and Sannikov (2012) provide a helpful overview of a large part of this literature. In particular our model builds on two distinct branches of financial macroeconomics. First, following Carlstrom and Fuerst (1997), Bernanke Gertler, and Gilchrist (1999) and others, we view lending to risky counterparties as a central element of the finance-macroeconomic nexus. For those authors the risky lending is to a distinct class of agents, entrepreneurs, who aim to maximize their utility, whilst for us the risky counterparties are the investment banks due to their investments in equities. These institutions have no objective other than profit maximisation as they are owned by all agents in the economy. Second, we draw on Clark (1984), Hancock (1985), Berger et al. (1987), Barth and Ramey (2001), Benk, Gillman and Kejak (2005), Goodfriend and McCallum (2007), and Ireland (2014) who propose or test distinct, but related, models of bank loan production. Neither of these branches of the literature incorporate default nor address the issue of the macroeconomic impact of differing vertical structures of financial intermediaries.[4]

The rest of the paper is set out as follows. Section 2 sets out the model. After describing the behavior of private agents and final goods producers it, the decisions of investment banks and their interaction with retail banks are analyzed including the optimal default decisions. Section 3 solves out for the general equilibrium of the baseline model and sets out how different assumptions concerning banking structure and bailouts affects the baseline model. Section 4 presents the social planning solution which facilitates a detailed analysis of the wedges of inefficiency in the decentralised economy. A numerical analysis of the model economy in Section 5 provides sensitivity analysis to welfare comparisons between universal and separated banking systems. Section 6 summarizes and concludes. Appendices contain additional calculations, derivations and proofs referred to in the text.

2 Macroeconomic Framework

The basic set up of the model is as follows: The economy consists of continua of households, monopolistically competitive, risk-neutral banks and final goods producers. There is also a government. Households consume the final goods, provide labour to the intermediary sector[5] and deposit their savings in the retail banks. The retail banks, if separate from

[4]The possibility of default of the depository institution has been considered in Angeloni and Faia (2013), and more recently in Gertler, Kiyotaki and Prestipino (2017). The overall focus of these papers is, however, rather different to ours. Two branches of banking services - retail and wholesale - are modelled in Gertler, Kiyotaki and Prestipino (2016). The overall impact of bailouts on optimal structure of financial intermediaries is not considered.

[5]Here, as in Gertler and Kiyotaki (2010), we assume that investment banks hold the equity of the intermediate producers. Thus, as equity holder, the investment bank determines the business strategy,

investment banks, lend to risky intermediary firms which use the funds to hire labour. The investment banks hold the equity of the risky entrepreneurs and make their hiring and production decisions before they observe their productivity and the demand for their output. That output is an intermediate good; that is, an input to the production of the final good. Investment banks have differing rates of profitability because they face idiosyncratic shocks. Because of idiosyncratic and common shocks, the values of banks' assets are stochastic, and some of the banks may default. The role of government is to bail out the banks when necessary and possible. The components of the model are now described in more detail.

2.1 Households

There is a continuum of identical households in the economy who evaluate their utility using the following criterion:

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\log(C_t) - \lambda N_t\right). \tag{1}$$

E_t denotes the expectations operator at time t, β is the discount factor, C_t is consumption and N_t is labour, λ is a time-invariant preference parameter.

In period t, agents have to decide how much of their current wealth to place in retail banks, D_t^h, given W_t, the nominal wage in period t, the expected return on deposits and Π_t, the corporate profits remitted to the individual. The household's budget constraint is

$$C_t + D_t^h = R_{t-1}^h \Gamma_t D_{t-1}^h + W_t N_t + \Pi_t. \tag{2}$$

Between date $t-1$ and the start of t deposit balances earn a nominal gross interest return of $\Gamma_t R_{t-1}^h$, where R_{t-1}^h is the gross interest each bank agrees to pay ex ante. However, the ex post return may be smaller if banks' assets at the end of the period are lower than $R_{t-1}^h D_{t-1}^h$. In that case banks will pay only proportion Γ_t^b of their obligations. If there is deposit insurance then Γ_t^g is provided by government. Therefore the proportion of the contracted return actually received by the depositors is $\Gamma_t = \Gamma_t^g + \Gamma_t^b$. If deposit insurance is not provided, $\Gamma_t^g = 0$. However, when deposit insurance is provided, there exists the possibility that profits may be so low in the economy that government may not have the capacity to bail out in full the depository institutions. The Γ_t reflects these eventualities, hence it is stochastic and $\Gamma_t \leq 1$. Thus Γ_t is an equilibrium object which will depend on the structure of the banking sector, government fiscal capacity and the confluence of shocks observed in each time period. We derive its form in different scenarios below.

including employment and the degree of risk taking.

Necessary conditions for an optimum include:

$$C_t = \frac{W_t}{\lambda}; \tag{3}$$

and

$$E_t \left\{ \Gamma_{t+1} R_t^h \frac{\beta C_t}{C_{t+1}} \right\} = 1. \tag{4}$$

2.2 The final goods sector

The production of final goods are common to all producers

$$Y_t = A_t X_t, \tag{5}$$

where X_t is an intermediate input, A_t may be thought of as an aggregate macro shock to output or as a utilization shock.

The production cost is $\frac{Q_t}{A_t} Y_t$, where Q_t is the real price of the output of the intermediate sector. It is straightforward to derive an aggregate real price, (6), and the aggregate demand for the financial product (7):

$$\frac{Q_t}{A_t} = \frac{1}{\mu^F}; \tag{6}$$

$$X_t = Y_t / A_t. \tag{7}$$

Here, μ^F is a monopolistic mark up in final goods production.

2.3 Banks

There are potentially two types of banks in the model, investment banks and retail banks. The output of investment banks, as noted, comprises a bundle of intermediary goods and services demanded by the final goods producers. Investment banks may be separated from retail banks. In that case, investment banks finance their activities by borrowing funds from retail banks. The only role for retail banks is to collect deposits from households and channel funds to the investment banks. In this loan market they are monopolistic competitors. When investment and retail banks are vertically integrated into a universal bank, there is no role for such a loan market.

The banking sector problems are now set out in detail.

2.4 The investment bank

Assume that investment and retail banks are separate. Agents deposit savings in retail banks. The retail banks bundle and sell these funds to an investment banking sector.

Each investment bank buys the entire equity of a single intermediate goods producer. The funds so invested, which were borrowed from the retail sector, pay the intermediate goods producers' wage bill ahead of selling their output to the final goods sector. So, one may think of the investment bank and the intermediate goods producer as one and the same entity which we do from here on, for simplicity.

Our investment banks are rather like the risky entrepreneurs in the models of Carlstrom and Fuerst (1997) and Bernanke Gertler, and Gilchrist (1999). As in Gertler and Kiyotaki (2010), banks here own a business which generates risky profits. Unlike them, however, we allow business risk to be sufficiently high that default on deposits is a real possibility. If output is below some value then these banks default, ending up with negative net assets, just like risky entrepreneurs. If these losses in aggregate are large, retail banks may not be able to repay depositors in full. The banks' losses may be made good by the taxpayer. If output is high enough, profit is remitted to private agents.

The investment bank cum intermediate firm produces output at $t+1$, $X_{t+1}(j)$, by employing labour at time t. Labour is homogeneous and is used with the following production technology to which all banks have access:

$$X_{t+1}(j) = \Omega_t \varepsilon_{t+1} e_{t+1}(j) N_t(j). \tag{8}$$

Here, $N_t(j)$ is the labour input employed by investment bank j, $\Omega_t > 0$ is the time t expected return common to all investment banks, ε_{t+1} is a shock that is also common to all banks and $e_{t+1}(j)$ is a $j-$specific shock. It is assumed that $e_{t+1}(j) \geq 0$, $\varepsilon_t \geq 0$, $E_t\varepsilon_{t+1} = 1$, and $E_t e_{t+1}(j) = 1$ and $(e_{t+1}(j), \varepsilon_{t+1})$ are independently distributed. The cumulative distributions of ε_{t+1} and e_{t+1} are denoted by $F^{\varepsilon}(\varepsilon)$ and $F^{e}(e)$, are time-invariant and common to all banks. (8) is undoubtedly a simple, yet very convenient, way to reflect risky activities; when we merge retail and investment banks later, the risk inherent in (8) will be incorporated onto each retail bank's balance sheet reducing risk diversification. In this way, we capture a key concern of policymakers that universal banks are more risky than separated banking structures.

At the start of period t the investment bank borrows the amount $B_t(j) = W_t N_t(j)$ from retail banks. In the next period the investment bank receives $Q_{t+1}(j)X_{t+1}(j)$, and pays $B_t(j)R_t^C$ to the retail bank, where $Q_{t+1}(j)$ denotes the price per unit $X_{t+1}(j)$, and R_t^C is the interest due on the loan.

The market for the output of the investment banking sector is assumed to be monopolistically competitive and the demand for output of bank j is

$$X_t(j) = \left(\frac{Q_t(j)}{Q_t}\right)^{-\eta} X_t. \tag{9}$$

Where $\eta > 1$ is the demand elasticity between the variety of products or the degree of competition[6]. The aggregate price next period, Q_{t+1}, and aggregate demand, X_{t+1}, are exogenous to the bank's decision. Combining (8) with (9) shows that the ex-post price depends on the realization of common and specific banking shocks $Q_{t+1}(j) = Q_{t+1}\left(\Omega_t \varepsilon_{t+1} e_{t+1}(j) \frac{N_t(j)}{X_{t+1}}\right)^{-1/\eta}$. So, the bank's assets at the end of period are

$$Q_{t+1}(j) X_{t+1}(j) = \left[\Omega_t \varepsilon_{t+1} e_{t+1}(j) N_t\right]^{1-1/\eta} X_{t+1}^{1/\eta} Q_{t+1}. \tag{10}$$

Expression (10) clearly shows that conditional expected profit depends not only on the productivity shocks, $\varepsilon_{t+1} e_{t+1}(j)$, but also on the state of the macroeconomic environment, represented by (X_{t+1}, Q_{t+1}).

2.5 Optimal default decision of investment banks

Ex-ante, the investment bank needs to decide on the level of borrowing/labour input. We suppose that investment banks have limited liability and act as though profit is bounded below at zero. So, assuming banks are risk-neutral, expected profit is

$$\begin{aligned} E_t \Pi_{t+1}(j) &= \max\left[E_t Q_{t+1}(j) X_{t+1}(j) - W_t N_t R_t^C, 0\right] \\ &= \max\left[E_t \left[\Omega_t \varepsilon_{t+1} e_{t+1}(j) N_t\right]^{1-1/\eta} X_{t+1}^{1/\eta} Q_{t+1} - W_t N_t R_t^C, 0\right]. \end{aligned} \tag{11}$$

The limited liability distortion means that banks will seek to maximize profits on a subset of states of nature. As a result, they will choose borrowing and a cut-off value for a composite of the shocks facing the bank below which default will occur. We now construct the bank's optimization problem in detail.

First, note that aggregate price and demand will depend on the realisation of both macro and banking shocks

$$Q_{t+1} = \frac{1}{\mu^F} A_{t+1}; \tag{12}$$

where we assume that $A_t = A_{t-1}^{\rho} u_t$, $0 < \rho < 1$ and u is lognormally distributed with $E(u) = 1$.

Since Ω and ε are common to all banks, the aggregate supply of intermediation is given by

$$X_{t+1} = \overline{N} \Omega_t \varepsilon_{t+1} \left[\int_0^{\infty} [e]^{\frac{\eta-1}{\eta}} \, dF^e(e)\right]^{\frac{\eta}{\eta-1}} = \overline{N} \Omega_t \varepsilon_{t+1} \Delta, \tag{13}$$

where $\overline{N}$ is the average number of employees at the other investment banks and $\Delta \equiv \left[\int_0^{\infty} [e]^{\frac{\eta-1}{\eta}} \, dF^e(e)\right]^{\frac{\eta}{\eta-1}}$ is the aggregate of idiosyncratic shocks across investment banks. There

[6] The aggregate demand for financial intermediation is defined over a basket of services indexed by j, $X_t \equiv \left[\int_0^1 X_t(j)^{\frac{\eta-1}{\eta}} di\right]^{\frac{\eta}{\eta-1}}$, where $\eta > 1$ is the elasticity of substitution. The aggregate price index is $Q_t = \left[\int_0^1 Q_t(i)^{1-\eta} di\right]^{\frac{1}{1-\eta}}$.

is no strategic interaction amongst the banks and $\overline{N}$ is treated as parametric by each bank. So, combining (12) and (13) means that (11) can be written, omitting time subscripts, as

$$\Pi|u\varepsilon e^{1-1/\eta} = N \max\left[\Omega u\varepsilon e^{1-1/\eta}\frac{1}{\mu^F}A^{\rho}\left(\frac{\overline{N}\Delta}{N}\right)^{1/\eta} - WR^C, 0\right]. \tag{14}$$

That expression is positive if and only if

$$u\varepsilon e^{1-1/\eta} > \varepsilon^D, \tag{15}$$

where

$$\varepsilon^D = \Lambda\left(\frac{\overline{N}}{N}\right)^{-1/\eta}, \tag{16}$$

and for purposes later on it is convenient to define

$$\Lambda = \frac{1}{\mu^F}\frac{W_t R^C}{\Omega_t A_t^{\rho}}(\Delta)^{-1/\eta}. \tag{17}$$

Formula (16) represents an ex-ante planned default threshold chosen by an individual bank taking macroeconomic factors, W_t, R_t^C, $\overline{N}_t$ and A_t, as given. However, the ex-post default rate depends on the realisation of the product of shocks, $s := u\varepsilon e^{1-1/\eta}$ where s is a random variable with lognormal density $f^s(s)$. If $s > \varepsilon^D$, then the bank will realize positive profits, otherwise profit are, in effect, zero. Hence, the complete investment banking problem can be written as:

$$\max_{N,\ \varepsilon^D} W_t R_t^C N_t \int_{\varepsilon^D}^{+\infty} \left[\left(\frac{s}{\varepsilon^D}\right) - 1\right] f^s(s)\, ds; \tag{18}$$

$$\text{s.t. } \varepsilon^D - \Lambda\left(\frac{N}{\overline{N}}\right)^{1/\eta} = 0. \tag{19}$$

In a symmetric equilibrium with $\frac{\overline{N}}{N} = 1$, so that (19) implies $\varepsilon^D = \Lambda$, where Λ is defined in

$$\int_{\Lambda}^{+\infty} \left[\frac{s}{\Lambda} - \frac{\eta}{\eta - 1}\right] f^s(s)\, ds = 0. \tag{20}$$

If the distribution of s is given, then Λ is just a constant which solves this integral equation. However, there may exist no, or many, solutions to this equation. In the appendix (Section 7.1) we define sufficient conditions for a solution to exist[7]. It is shown that a lognormal distribution satisfies those conditions and provides a unique Λ as a solution. Moreover, under a lognormal distribution the second order conditions for an optimum are satisfied. Importantly, one observes that this equation has implications for the relationship between the degree of competition and equilibrium default to which we turn presently.

Given Λ one may compute the equilibrium wage

$$W_t = \frac{1}{\mu^F}\Omega_t A_t^{\rho}\frac{\Lambda}{R_t^C}(\Delta)^{1/\eta}. \tag{21}$$

[7] It is interesting to note that Bernanke, Gertler and Gilchrist (1999) derive the same condition.

This relation implies that the equilibrium wage increases in the tolerance to risk, Λ, of the investment banks. Or, to state this last observation differently, banks which are prepared to take on higher risks will offer higher wages.

Finally, one can compute the default threshold, conditional on the realisation of common shock from (15)

$$e^D = \left(\frac{\Lambda}{u\varepsilon}\right)^{\frac{\eta}{\eta-1}}. \tag{22}$$

Every bank with an idiosyncratic shock lower than e^D will be in default while those for whom $e \geq e^D$ will be able to meet their commitments. From (22) it follows that the conditional probability of default depends only on the unexpected macro shock u and the realisation of the common shock to the investment banking sector, ε. At an optimum, it does not depend on the expected gross return to the sector, Ω, nor on the current state of the macroeconomy, A_t. However it depends on the distribution of the aggregate business shock f^s as emphasized by Christiano, Motto and Rostagno (2014).

We turn now to the retail banking sector.

2.6 Retail bank and the credit spread

There is a continuum of risk-neutral, retail banks indexed by i. Banks pay a state dependent, contractual interest rate on deposits of R_t^h, if possible. That deposit rate will be common across banks and need not be indexed by i. In the loans market, banks are monopolistic competitors and set loan rates, $R_t^c(i)$. So, following Aksoy et al. (2012), banks face the following demand for loans

$$B_t^c(i) = \left(\frac{R_t^c(i)}{R_t^c}\right)^{-\delta} B_t^C. \tag{23}$$

Here $B_t^c(i)$ is bank i's lending, R_t^C is a measure of the average interest rate on loans, $R_t^C = \left[\int_0^1 R_t^c(i)^{1+\delta} di\right]^{\frac{1}{1+\delta}}$, B_t^C is aggregate demand for loans, $B_t^C = \left[\int_0^1 B_t^c(i)^{\frac{\delta-1}{\delta}} di\right]^{\frac{\delta}{\delta-1}}$, and $\delta > 1$ is the elasticity of substitution between loans. The objective of each bank, therefore, is to maximize expected profits by choosing the rate charged on lending. If all borrowers remain solvent, the retail bank will earn nominal return R_t^C per unit loaned. In the case of default, the assets of the borrower are repossessed pro rata by the retail bank.

Retail banks maximize expected profit, $E_t\Psi_{t+1}$, given the demand for loans, (23), and knowing that their liabilities are limited:

$$E_t\Psi_{t+1}(R_t^c(i)) = E_t \max(R_t^c(i)B_t^c(i) - R_t^h B_t^c(i), 0). \tag{24}$$

Although the profits of individual retail banks are in effect bounded below by zero, the net worth of the retail bank sector is ultimately determined by outturns in the investment banking sector. In some states, an investment bank may not be able to repay its loan in full. One may characterize as follows that portion of the loan which can be repaid by the investment bank. At period $t+1$ every investment bank j has given liability, $W_t R_t^C N_t$, whilst its assets are stochastic and equal to $\Omega u\varepsilon e(j)^{1-1/\eta}\frac{1}{\mu^F}A_t^\rho N_t$. Thus, the assets to liabilities ratio can be written as

$$\frac{\Omega u\varepsilon e^{1-1/\eta}A_t^\rho \Delta^{1/\eta}}{WR^C\mu^F}=\frac{u\varepsilon e^{1-1/\eta}}{\Lambda}.$$

Therefore the borrower is in default if $\frac{u\varepsilon e^{1-1/\eta}}{\Lambda}<1$, and the gross return generated by the borrower in default will be $\frac{u\varepsilon e^{1-1/\eta}}{\Lambda}R_t^c(i)$. Let $\Gamma_{u\varepsilon e}^b(u\varepsilon e)$ be the ratio of actual to contractual return conditional on a particular realisation of the shocks. That is,

$$\Gamma_{u\varepsilon e}^b(u\varepsilon e)=\min\left(\frac{u\varepsilon e^{1-1/\eta}}{\Lambda},1\right).$$

After averaging over all possible idiosyncratic shocks, one obtains the expected ratio of actual to contractual return conditional on the realisation of macro and systemic shocks, $u\varepsilon$:

$$\Gamma^b(u\varepsilon)=u\varepsilon\int_0^{\left(\frac{\Lambda}{u\varepsilon}\right)^{\frac{\eta}{\eta-1}}}\frac{e^{1-1/\eta}}{\Lambda}f^e(e)de+1-F^e\left(\left(\frac{\Lambda}{u\varepsilon}\right)^{\frac{\eta}{\eta-1}}\right). \tag{25}$$

In formula (25) the first term is expected revenue from repossession of the assets of the investment banks which are in default. The remaining term is the expected revenue from non-defaulting banks. It is useful to establish some basic properties of the function $\Gamma^b(u\varepsilon)$. We do this in:

Proposition 1 $\Gamma^b(u\varepsilon)$ *is an increasing and concave function. Therefore, more favourable systemic or macro shocks bring average returns closer to the contractual return. Moreover* $\Gamma^b(0)=0$ *and* $\lim_{x\to\infty}\Gamma^b(x)=1$.

Proof. *See Appendix* ■

The profit of the retail bank conditional on the realisation of the aggregate shocks will be

$$\Psi_{t+1}(R_t^c(i),u\varepsilon)=\max\left[\frac{R_t^c(i)}{R_t^h}\Gamma^b(u\varepsilon)-1,0\right]B_t^c(i)R_t^h. \tag{26}$$

Using the demand for loans (23) one writes

$$\Psi_{t+1}(R_t^c(i),u\varepsilon)=\max\left[\frac{R_t^c(i)}{R_t^h}\Gamma^b(u\varepsilon)-1,0\right]\left(\frac{R_t^c(i)}{R_t^c}\right)^{-\delta}B_t^C R_t^h. \tag{27}$$

For any $R_t^c(i)$ there is a value of common shock $u\varepsilon=y$ below which retail banks will default on their obligations. Therefore, the expected profit maximization problem (24) can be written as

$$\begin{aligned}\max_{R_t^c(i),y} E\Psi_{t+1} &= \left[\int_y^{+\infty}\left(\frac{R_t^c(i)}{R_t^h}\Gamma^b(x)-1\right)f_{u\varepsilon}(x)dx\right]\left(\frac{R_t^c(i)}{R_t^c}\right)^{-\delta}B_t^C R_t^h \\ \text{s.t.}\quad \frac{R_t^c(i)}{R_t^h}\Gamma^b(y) &= 1.\end{aligned} \tag{28}$$

The first order conditions imply that in a symmetric equilibrium the retail banks default threshold y is defined as

$$\int_{y}^{+\infty}\left(\frac{\Gamma^{b}(x)}{\Gamma^{b}(y)}-\frac{\delta}{\delta-1}\right)f_{u\varepsilon}(x)dx=0, \tag{29}$$

and the spread R_t^C/R_t^h in equilibrium is

$$\frac{R_t^C}{R_t^h}=\frac{1}{\Gamma^{b}(y)}. \tag{30}$$

Proposition 2 *There exists a solution to (29) which also satisfies the second order conditions.*

Proof. *See the appendix.* ■

It is interesting to note that the wedge (credit spread) imposed by separated banking is larger than the familiar monopolistic wedge:

Proposition 3 *The mark up charged by retail banks is greater than the standard monopoly pricing wedge.*

Proof. From (29) and (30) it follows that

$$\frac{R_t^C}{R_t^h}=1/\Gamma^{b}(y)=\frac{\delta}{\delta-1}\frac{[1-F_{u\varepsilon}(y)]}{\int_{y}^{+\infty}\Gamma^{b}(x)f_{u\varepsilon}(x)dx.}>\frac{\delta}{\delta-1}. \tag{31}$$

■

Since $\Gamma^{b}(x)\leq 1$, it follows that $\int_{y}^{+\infty}\Gamma^{b}(x)f_{u\varepsilon}(x)dx<[1-F_{u\varepsilon}(y)]$. The ratio $\mu^R=\frac{[1-F_{u\varepsilon}(y)]}{\int_{y}^{+\infty}\Gamma^{b}(x)f_{u\varepsilon}(x)dx.}>1$ represents the contribution of risk to the mark up (which is not related to market concentration). If returns were certain ($\Gamma^b=1$) then $\mu^R=1$ and only monopoly power would cause inefficiency. Therefore, expression (31) shows that uncertainty in the return on retail lending makes the wedge in the banking sector larger than it otherwise would be.

2.7 Deposit insurance

It is now possible to characterize in more detail the government's behavior. In this stylized model government's only function is to raise funds for deposit insurance. The size of the government bail out is denoted G. It is assumed that government intervention is costly. Such costs, denoted here by $g(G_t)$, are generally associated with monitoring costs and distortive taxation. For tractability, we assume this g function is linear in G. It is then the case that

$$Y_t = C_t + gG_t, \quad g \geq 0. \tag{32}$$

As one knows how profitable lending is typically likely to be, one may now compute the likely size of government intervention. First note that the actual (or average) return, R_t^{ca} per unit borrowed by investment banks, will depend on the realisation of the macro environment,

$$R_t^{ca}(u\varepsilon) = R^C \Gamma^b(u\varepsilon). \tag{33}$$

Government intervention will occur if $R_t^{ca}(u\varepsilon) < R^h$, that is, when $u\varepsilon < y$. The size of the government bailout, G, depends on a number of considerations as follows. To begin with, if the banking firm is separated (indicated by the S superscript as before), then for one dollar of deposits, agents will receive *from the bank* the interest rate R_t^h times $\Gamma_{u\varepsilon}^{bS}(u_{t+1}\varepsilon_{t+1})$, where

$$\Gamma^{bS}(u_{t+1}\varepsilon_{t+1}) = \min\left(\frac{\Gamma^b(u\varepsilon)}{\Gamma^b(y)}, 1\right). \tag{34}$$

Function $\Gamma^{bS}(u_{t+1}\varepsilon_{t+1})$ represents the proportion of contracted deposits which depositors will obtain from retail banks in the case where deposits are not insured by the government.

It also follows that for any dollar of bank liabilities, the required bailout or deposit insurance is $1 - \Gamma^{bS}(u\varepsilon) \geq 0$. However, the economy-wide budget constraint (32) imposes a natural restriction on the size of G_t. For example, in the event that total national output is required to repay households under deposit insurance, the return on deposits is the only source of funds for consumption, and equation (32) becomes $Y_t = (1 + g)G_t$. And whilst it is a rare event in our model, it may even happen that total insurance exceeds fiscal capacity. Therefore the size of government intervention is restricted by the size of GDP, $G_t \leq Y_t/(1+g)$. In sum then, we assume that government guarantees the following compensation to the public

$$G_{t+1}^{gS} = N_t W_t R_t^h \min\left(\frac{Y_{t+1}}{N_t W_t R_t^h} \frac{1}{1+g}; \left(1 - \Gamma^{bS}(u_{t+1}\varepsilon_{t+1})\right)\right), \tag{35}$$

where the first term after the min operator reflects the fact that the required bailout might exceed fiscal capacity. In that case, the government is unable to fulfil its desire to see returned R_t^h per dollar deposited. Using the expressions for aggregate output, financial services, wages and interest spread, that is (5), (13), (21) and (30), it follows that

$$G_{t+1}^{gS} = N_t W_t R_t^h \min\left(\mu^F \frac{u_{t+1}\varepsilon_{t+1}}{\Gamma^b(y)\Lambda\left(\Delta\right)^{1/\eta-1}} \frac{1}{1+g}; \left(1 - \Gamma^{bS}(u_{t+1}\varepsilon_{t+1})\right)\right). \tag{36}$$

Next, define the ratio of government insurance to deposits as

$$\Gamma_{u\varepsilon}^{gS}(u_{t+1}\varepsilon_{t+1}) = \min\left(\frac{g}{\widetilde{g}} \frac{1}{\Gamma^b(y)} \frac{u_{t+1}\varepsilon_{t+1}}{1+g}; \left(1 - \Gamma^{bS}(u_{t+1}\varepsilon_{t+1})\right)\right).$$

where $\widetilde{g} := g\frac{1}{\mu^F}\left(\Delta\right)^{1/\eta-1}\Lambda$. Finally, the return to depositors equals the sum received from the banks and from the government: $\Gamma^{SI} = \left(\Gamma^{gS}(u_{t+1}\varepsilon_{t+1}) + \Gamma^{bS}(u_{t+1}\varepsilon_{t+1})\right)$,[8] so that the

[8] That is: $\Gamma_{u\varepsilon}^{gS}(u_{t+1}\varepsilon_{t+1}) + \Gamma_{u\varepsilon}^{bS}(u_{t+1}\varepsilon_{t+1}) = \min\left(\frac{g}{\widetilde{g}}\frac{1}{\Gamma^b(y)}\frac{1}{1+g} + \Gamma_{u\varepsilon}^{bS}(u_{t+1}\varepsilon_{t+1}); 1\right)$.

consumption Euler equation takes the following form $E_t\left\{\frac{\beta C_t}{C_{t+1}}R_t^h\Gamma^{SI}\right\} = 1$. Superscript "$S$" indicates "$S$eparated banking with Insurance ".

It is useful to note that the likelihood and size of government support to the banking system declines in the credit spread. To see this more explicitly, consider the case where the government faces no fiscal capacity constraints. In that case government intervention would be $G_{t+1} = W_t N_t R_t^h \max\left[\left(1 - \frac{R_t^C}{R_t^h}\Gamma^b(u\varepsilon)\right), 0\right]$and the negative effect of the spread is readily apparent.

3 Financial structure and economic outcome

Clearly, the precise form of the equilibrium relations of the model change depending on the nature of the shocks, whether or not there is universal banking and deposit insurance and government fiscal capacity. We will contrast four variants of the model with different financial structures indexed by superscript $J = \{UI,\ UN,\ SI,\ SN\}$, where S stands for separated, U for universal, I indicates that the government provides deposit insurance and N means that there is no deposit insurance. The variants of the model will differ along three key dimensions: (i) the return on deposits, which is reflected in the Euler equation; (ii) the size of government intervention and (iii) the credit spread.

Table 1 below summarizes the differences where, as before, Γ^{bJ} is the proportion of the deposits liabilities paid by the banks, and Γ^{gJ} is the proportion which is paid for by the government. The safety of deposit holding Γ^{bJ} increases with spread and here we can see a partial equilibrium benefits of separated banking. In case when deposits are protected, the size of government intervention is smaller. So credit spread reduces the risk of the default of the depositary institution and reduces the cost of banks' bailouts.

Table 1: Difference across the models

$$\mathrm{sp}^J = \left\{\begin{array}{ll} 1; & \text{for Universal banking} \\ 1/\Gamma^b(y); & \text{for Separated banking,} \end{array}\right\};$$

$$\Gamma^{bJ}(u\varepsilon) = \min\left(\Gamma^b(u\varepsilon)\mathrm{sp}^J, 1\right);$$

$$\Gamma^{gJ}(u\varepsilon) = \left\{\begin{array}{ll} \min\left[\frac{g}{g}\frac{\mathrm{sp}^J}{1+g}; \left(1 - \Gamma^J_{u\varepsilon}(u\varepsilon)\right)\right] & \text{with insurance;} \\ 0, & \text{without insurance;} \end{array}\right\}$$

The safety of deposit holding Γ^J increases with spread and here we can see a partial equilibrium benefits of separated banking: When deposits are covered by deposit insurance, the size of government intervention is smaller. So, the credit spread reduces the risk of the default of the depositary institution and reduces the cost of banks' bailouts. However, the same spread increases the marginal cost of the intermediary sector. Therefore, a general equilibrium model is required to analyse that trade off between risk and the size of the economy.

The generic equations of the model are set out below as part of the definition of the decentralised equilibrium:

A decentralised equilibrium with separated banking and bailouts is a set of plans, $\{C_{t+k},\, Y_{t+k},\, N_{t+k},\, W_{t+k},\, X_{t+k,}\, Q_{t+k},\, R^c_{t+k},\, R^h_{t+k},\}_{k=0}^{\infty}$, given initial conditions, $\{A_{t-1}, N_{t-1}, R^h_{t-1}, R^c_{t-1}, W_{t-1}\}$, and exogenous shocks, $\{u_{t+k},\ \varepsilon_{t+k},\ \Delta_{t+k}\}_{k=0}^{\infty}$, and satisfying conditions (M1)-(M10).

Table 2:	Model Equations	
Euler Equation	$\frac{1}{\beta C_t} = E_t \left\{ \left[\Gamma^{gJ}_{u\varepsilon}(u_{t+1}\varepsilon_{t+1}) + \Gamma^{bJ}_{u\varepsilon}(u_{t+1}\varepsilon_{t+1}) \right] R^h_t \frac{1}{C_{t+1}} \right\}$	(M1)
Deposit insurance	$G_{t+1} = N_t W_t R^h_t \Gamma^{gJ}_{u\varepsilon}(u_{t+1}\varepsilon_{t+1})$	(M2)
Credit spread	$\frac{R^C_t}{R^h_t} = \mathrm{sp}^J$	(M3)
Wage Setting	$W_t = \lambda C_t$	(M4)
Labour demand	$W_t = \frac{1}{\mu^F} \Omega_t A^\rho_t \frac{\Lambda}{R^C} (\Delta)^{1/\eta}$	(M5)
TFP	$A_{t+1} = A^\rho_t u_{t+1}$	(M6)
FI demand	$Q_t = A_t \frac{1}{\mu^F}$	(M7)
Final good production	$Y_t = A_t X_t;$	(M8)
FI supply	$X_{t+1} = \Omega \Delta \varepsilon_{t+1} N_t$	(M9)
Resource constraint	$C_t = Y_t - g(G_t)$	(M10)

One observes that equations (M4-M10) are identical across the variants of the model, whereas possible differences emerge via the anticipated return on deposits, (M1), the size of government intervention, (M2), and the credit spread, (M3).

The above bloc of equations can be used to derive tractable, closed-form expressions for equilibrium consumption, labour and the deposit rate. These will be used in subsequent sections of the paper to analyze welfare under different scenarios; that is, with and without insurance and under separated and universal banking. Appendix 7.6 provides the details of the closed-form solution of the model. What is key, is solving in the first instance for the labour input:

$$N^J = \frac{\beta}{\lambda} \frac{1}{\mu^F} \Lambda \left(\Delta\right)^{1/\eta - 1} \Upsilon^J. \tag{37}$$

In this expression, equilibrium labour is seen to be a function of agents' preference parameters in the usual way, demand factors (reflected in the markup), the default cut-off value and the aggregate of idiosyncratic shocks across investment banks. Finally, it is a function of the variable Υ^J which in turn depends on financial structure. Table 3 expands on the definition of that final factor:

Table 3: Financial structure and Employment		
	Universal (U)	Separated (S)
I	$\Upsilon^{UI} = \int\limits_0^{+\infty} \frac{\Gamma^g(u\varepsilon)+\Gamma^b(u\varepsilon)dF_{u\varepsilon}}{u\varepsilon-\widetilde{g}\Gamma^g(u\varepsilon)}$	$\Upsilon^{SI} = \frac{R_t^h}{R_t^C} \int\limits_0^{+\infty} \frac{\left[\Gamma^{gS}_{u\varepsilon}(u\varepsilon)+\Gamma^{bS}_{u\varepsilon}(u\varepsilon)\right]dF_{u\varepsilon}}{u\varepsilon-\widetilde{g}\Gamma^{gS}_{u\varepsilon}(u\varepsilon)\Gamma_{u\varepsilon}(y)}$
N	$\Upsilon^{UN} = \int\limits_0^{+\infty} \frac{\Gamma^b(u\varepsilon)dF_{u\varepsilon}}{u\varepsilon}$	$\Upsilon^{SN} = \frac{R_t^h}{R_t^C} \int\limits_0^{+\infty} \frac{\Gamma^{bS}_{u\varepsilon}(u\varepsilon)dF_{u\varepsilon}}{u\varepsilon}$

The relationships in Table 3 will prove central in establishing a number of properties concerning the link between equilibrium employment and the financial structure of the economy. First we look at employment and output and then briefly at interest rates. Note that production in the economy will have the same relation to Υ as does N as it positively depends on labour input. Moreover, given these relations, one can immediately establish the impact of deposit insurance on output and employment.

We turn now to deposit insurance.

3.1 Deposit insurance and the labour market

There is a positive effect of deposit insurance on labour and output which we establish in the following proposition:

Proposition 4 *Equilibrium employment and output are larger in an economy with government deposit insurance.*

Proof. One can easily show that $\Upsilon^{UI} > \Upsilon^{UN}$. Note that $\Gamma^g(u\varepsilon) + \Gamma^b(u\varepsilon) > \Gamma^b(u\varepsilon)$ and $\frac{1}{u\varepsilon-\widetilde{g}\Gamma^g(u\varepsilon)} > \frac{1}{u\varepsilon}$.

Similarly, $\Upsilon^{SI} > \Upsilon^{SN}$, as $\Gamma^{gS}(u\varepsilon) + \Gamma^{bS}(u\varepsilon) > \Gamma^{bS}(u\varepsilon)$ and $\frac{1}{u\varepsilon-\widetilde{g}\Gamma^{gS}_{u\varepsilon}(u\varepsilon)\Gamma_{u\varepsilon}(y)} > \frac{1}{u\varepsilon}$. ■

Proposition 4 states that deposit insurance increases employment. However, it is important to note that deposit insurance impacts the equilibrium relations of the economy through a number of channels.

In particular, there are two channels which we now briefly describe and which we analyse further in subsequent sections on welfare and efficiency of equilibrium. First, deposit insurance $\Gamma^{gS}(u\varepsilon)$ works to increase the supply of labour and deposits. That is because it stabilizes the return to savings; by implication the return to working is more certain. Second, from (M10), g is seen to have a direct negative effect on C_{t+1}. Smaller expected future consumption causes an increase in savings and therefore supply of deposits. Moreover, given the cost of government intervention one has to work more to maintain a given level of consumption.

To derive these effects explicitly consider the Euler equation

$$1 = E_t \frac{u'(C_{t+1})}{u'(C_t)} \beta R_t^h \Gamma,$$

where the labour supply is $u_c'(C_t) = v_n'/W$, and v_n' is the marginal disutility from labour. In the model utility is separable in consumption and labour, and v_n' is constant so that

$$1 = \frac{\beta W R_t^h}{v_n'} E_t \Gamma u_c'(C_{t+1}).$$

As C_{t+1} depends on N_t, and some parameter g, such that $C_{t+1} = C(N_t, g)$, where $C_N > 0$, $C_g \leq 0$, one may write

$$v_n'(N_t) = \beta W R_t^h E_t \Gamma u'(C_{t+1}(N_t, g)). \tag{38}$$

It follows by the implicit function theorem that

$$\frac{dN}{dg} = -\frac{\beta W R_t^h E_t \Gamma u''(C(N,g))C_g}{\beta W R_t^h E_t \Gamma u''(C(N,g))C_N - v''(N)} > 0. \tag{39}$$

In establishing that sign, it is assumed that $v''(N) > 0$, $u''(C) < 0$, and that

$$\frac{dN}{d\Gamma} = -\frac{\beta W R_t^h E_t u'(C(N,g))}{\beta W R_t^h E_t \Gamma u''(C(N,g))C_N - v''(N)} > 0. \tag{40}$$

Therefore, labour supply increases in the degree of certainty of the deposit return, Γ, and the inefficiency of government, g. The latter effect is due to a precautionary motive similar to that discussed in Kimball (1990).

3.2 Vertical integration and the labour market

It is also interesting to observe that employment, and therefore output, is greater in the model with universal banking:

Proposition 5 *Universal banking results in higher employment in equilibrium than a separated banking system.*

Proof. *One can easily show that $\Upsilon^{UN} > \Upsilon^{SN}$ since, by definition, $\Gamma^{bS}(u\varepsilon) = \min\left(\frac{\Gamma^b(u\varepsilon)}{\Gamma^b(y)}, 1\right)$ and therefore $\Gamma^b(y)\Gamma^{bS}(u\varepsilon) = \min\left(\Gamma^b(u\varepsilon), \Gamma^b(y)\right) \leq \Gamma^b(u\varepsilon)$. Similarly, $\Upsilon^{UI} > \Upsilon^{SI}$, which we show in the appendix 7.7. Reference to formula (37) completes the proof.* $\blacksquare$

It is not very surprising that vertical integration eliminates a credit spread and promotes production. Proposition 5 shows exactly this. However, it can only be welfare improving if production is below the optimal level because of some other distortions. Moreover, larger output does not guarantee larger consumption since the cost of deposit insurance may increase more than output. It is shown below that when deposit insurance is provided, the cost of bailouts is larger with a universal banking structure. Therefore it is possible that when the economy is hit by an adverse shock, consumption is smaller under universal banking due to the high cost of government intervention.

4 Efficiency, welfare and financial structure

In this section we summarise all the wedges of inefficiency in our model economy. We will derive the social planner's solution and then analyse the wedges created by each agent in decentralised equilibrium.

4.1 Social planner problem

A general version of the social planner's problem relevant to the present set-up is as follows:

$$U(N) = \max_{N} E_0 \sum_{t=0}^{\infty} \beta^t \left(u(C_t) - v(N_t) \right) \tag{41}$$

with the following constraints imposed by production technology

$$\begin{aligned} Y_{t+1} &\leq F(N_t, u_{t+1}) \\ C_{t+1} &\leq Y_{t+1}. \end{aligned}$$

The optimal choice is given by

$$\beta E_t u_c(F(N_t, u_{t+1})) F_N(N_t, u_{t+1}) = v'(N_t).$$

The left hand side of this expression reflects the marginal benefit of an extra unit of labour, and the right hand side reflects the cost. This is taken as the benchmark against which the decentralized outcomes under different scenarios for banking and bailouts are compared. In the model set out above the problem is $\max E_0 \sum_{t=0}^{\infty} \beta^t \left(\log(C_t) - \lambda N_t \right)$, subject to $C_t = A_t \Omega \Delta \varepsilon_t N_{t-1}$. The labour supply optimizing this objective is implied by $\beta E_t u_c(F(N_t, u_{t+1})) F_N(N_t, u_{t+1}) = v'(N_t)$ and is easily seen to be $N^* = \beta/\lambda$. One can compare that expression with decentralised solution (37).

4.2 Decentralised structure of the economy

It will be useful to set out carefully all the distortions attendant with the decentralised equilibrium. We will present it in form of

$$\beta E_t u_c(F(N_t, u_{t+1})) F_N(N_t, u_{t+1}) = \mu v'(N_t).$$

Here, $\mu > 1$ will present the total inefficiency wedge, which can be decomposed vertically into five different wedges accounting for each economic agent and representing their marginal benefits to marginal costs ratio:

$$\mu = (\mu^C \times \mu^I \times \mu^F) \times \mu^H \times \mu^g.$$

Here $\mu^F > 0$ is the monopolistic mark up in final goods production. The other wedges are more interesting.

First we consider the wedges of the banking sector. Retail banks take deposits and produce loans. Therefore the return on deposits, R^h, may be considered as the marginal cost facing retail banks whilst the return on credit, R^C, is the marginal benefit (profit). It is therefore straightforward to compute the wedges between costs and benefits for retail banking, $\mu^C = R_t^C / R_t^h$, and vertical integration would simply eliminate this margin.

As for investment banks, their marginal cost is given by $R_t^C W_t$, and $Q_{t+1}\Omega\Delta\varepsilon_{t+1}$ is their marginal profit. The mark up of investment banking is therefore $\mu^I = E_t \frac{Q_{t+1}\Omega\Delta\varepsilon_{t+1}}{R_t^C W_t}$. Combining this with (21), one finds that

$$\mu^I = (\Delta)^{1-1/\eta} / \Lambda, \tag{42}$$

which declines in the planned default threshold Λ. Therefore, there is a clear trade off between risk and inefficiency: the larger is the default rate, the smaller is the wedge imposed by the risky financial intermediary. It is also easy to show that μ^I is smaller than the monopolistic mark up $\frac{\eta}{\eta-1}$ due to the limited liability assumption.

The decentralised solution in (37) may be presented as $N_t^J = N^* \frac{1}{\mu^I} \frac{1}{\mu^F} \Upsilon^J$, where Υ^J is the product of three wedges imposed by households, retail banking (if it is separated from the investment bank) and government.

We will refer to the product of banking and production sector wedges as the "production wedge" and denote it $\mu^m := \mu^C \times \mu^I \times \mu^F$. The marginal cost of the integrated production line is $R_t^H W_t$ and the expected marginal benefit is $E_t F_N(N_t)$. So we can see that

$$E_t F_N(N_t) = \mu^m \times R_t^H W_t. \tag{43}$$

For the case when government intervention is costless, $Y_{t+1} = C_{t+1}$, we formally define the "household wedge" as the residual after the "production wedge" is accounted for. Then, the following equality will hold

$$\beta E_t u_c(F(N_t, u_{t+1})) F_N(N_t, u_{t+1}) = \mu^H \times \mu^m \times v'(N_t).$$

Combining that with (43) gives

$$W_t R_t^h \beta E_t u_c(F(N_t, u_{t+1})) F_N(N_t, u_{t+1}) = \mu^H v'(N_t) E_t F_N(N_t). \tag{44}$$

Finally, to gain an expression for the household wedge, we use that expression along with the labour supply and the Euler equation for saving as in (38), to get[9],

$$\mu^H = \frac{E_t\left[u'(F(N_t, u_{t+1}))F'(N_t)\right]}{E_t F'(N_t) E_t\left(\Gamma^J u'(F(N_t, u_{t+1}))\right)}.$$

If there is uncertainty about savings, then $E_t\left(\Gamma^J\right) < 1$. Deposit protection increases Γ^J and therefore reduces the households wedge. This provides another reason, in addition to avoiding banks runs, why deposit insurance may be beneficial.

However, when deposit protection is costly to deliver we need to account for an extra wedge. That causes output and consumption to diverge, and increases in the size of government intervention. Consequently, deposit insurance reduces the households wedge but increases the government intervention wedge and the overall net benefit depends on how costly government intervention is.

5 Welfare effects of universal and separated banking

In this section we summarise the economic factors in our model which make separated banking socially preferred to vertically integrated banking. This exercise may help us understand why some countries have evolved more separated banking systems while others have adopted more vertically integrated structures between deposit taking and risky investment. The model that has been developed is capable of assessing welfare effects associated with universal and separated banking in the face of different types of shocks. For instance, it has been demonstrated that labour supply is larger under universal banking. That is because the double marginalization problem is avoided boosting the demand for labour. We have also shown that costless deposit insurance increases the labour supply and moves the economy towards its efficient level of output. However, when deposit insurance is costly, the production to consumption wedge may be relatively large. In that case, separated banking with bailouts may be closer to the efficient outturn compared with universal banking with bailouts.

Another perspective on comparing welfare across banking structures is to enquire how likely are policymakers to regret adopting universal versus separated banking. This is a potentially interesting question as policy in this area is likely to be held fixed for extended periods. That was case, for example, with the split between retail and investment

[9] In an economy with safe deposits, $\Gamma^J = 1$, μ^H offsets the production wedge, because $\mu^H = \frac{E_t\left[u'(C_{t+1})F'(N_t)\right]}{E_t F'(N_t) E_t(u'(C_{t+1}))} = 1 + \frac{cov(u'(C_{t+1})F'(N_t))}{E_t F'(N_t) E_t(u'(C_{t+1}))} < 1$, as positive productivity shock increases F' and reduces $u'(C_{t+1})$, so the covariance term is negative.

banking in the US for much of the twentieth century. So it is interesting to ask what the model implies about *ex post* or average welfare under differing market structures for different realisations of the shocks. For example, it may be that a decision to permit or prohibit universal banking is welfare decreasing ex ante, but welfare increasing on average. Specifically, if banks are broken up (universal banks are prohibited) at the start of the period and a positive common shock is realized, then in principle one might have preferred in this state of nature to have had a universal banking structure. So, the average welfare comparison of different banking structures and the likelihood of regretting having a particular banking structure are both of interest. The welfare assessment of these various possibilities turns on certain key factors in the present model such as how volatile are common and idiosyncratic shocks; how competitive are retail and investment banking sectors; how generous is the deposit insurance scheme and how distortive is government intervention. We turn now to some examples briefly to illustrate these possibilities.

5.1 Comparative benefits of separated banking

The model is as yet too simple to be the basis for any serious policy advice but may be useful in indicating how certain key attributes of the economic environment interact and sway judgement on banking structure. We take the following parameterization as our starting point: $\sigma_e = 0.1$, $\sigma_u = 0.1$, $\sigma_\varepsilon = 0.1$, $\eta = 4$, $\mu^F = 1.14$, $\delta = 10$. For our bench mark simulation we make an assumption that the excess cost of fund raising is 20% of the value ($g = 0.2$), which is consistent with some estimates reported in Allgood et al (1998). Under this parameterization the expected default rate of investment banks is 5%; ($F_s(\Lambda) = 5\%$). That corresponds to an expected default rate for the retail banks of 0.5% ($F_{ue}(y) = 0.5\%$), implying that retail banks rarely fail. In fact, it implies a government bailout once in every 200 years. Notation $E\left(W^{UI} - W^{SI}\right)$ represents the expected relative welfare gain of universal banking as compare to a separated banking system measured in consumption equivalent. The probability of an adverse common shock which makes universal banking less desirable than separated banking is denoted by $\Pr(W^{UI} < W^{SI})$. Below we conduct some sensitivity analysis to investigate how different economic factors affect the welfare gains from the vertical integration.

5.1.1 Government efficiency

Government efficiency in bank bailout turns out to be one of the more significant drivers in the welfare analysis. More costly government intervention reduces consumption, increases labour supply and therefore reduces welfare. We investigated the desirability of universal banking for various degrees of inefficiency in terms of government action. The data are summarized in the following table.

Table 4. Government efficiency

	$g = 0.1$	$g = 0.2$	$g = 0.4$
N^{UI}	0.6533	0.6535	0.6541
N^{SI}	0.5861	0.5861	0.5862
$E\left(W^{UI} - W^{SI}\right)$	4.16%	4.14%	4.11%
$\Pr(W^{UI} < W^{SI})$	0	0	0.2%
$F_s(\Lambda)$	5.3%	5.3%	5.3%
$F_{u\varepsilon}(y)$	0.47%	0.47%	0.47%

The conclusion from the simulation is that universal banking is more likely preferable the smaller is the cost of the bank bail out, g. Whilst bailouts are larger and more frequent under universal banking, universal banking eliminates a double margin wedge. Taken together, the benefits from fewer wedges and relatively efficient government intervention may outweigh the costs of larger bailouts. Simulations suggest that for separated banking to be preferable, the composite shock has to be very bad indeed and government intervention highly distortive; in that case $E(W^{UI} - W^{SI}) < 0$ indicating that agents would require compensation on average to be indifferent between universal and separated banking. Indeed, the likelihood of universal banking being dominated, even under the most adverse of circumstances, is vanishingly small, $\Pr(W^{UI} - W^{SI} < 0) = 0.2\%$.

5.1.2 Idiosyncratic volatility

Christiano, Motto and Rostagno (2014) suggest that shocks to the volatility of cross-sectional idiosyncratic uncertainty are important in explaining the business cycle[10]. Indeed, in our model higher volatility results in a larger credit spread, which is a major distortion in the model with separated banking. Therefore, elimination of that distortion is likely to be welfare enhancing, ceteris paribus. Our simulation results in Table 5 confirm that intuition.

[10] See also De Fiore, Teles and Tristani, (2011).

Table 5. Idiosyncratic volatility

$g = 0.2$	$\sigma_e = 0.1$	$\sigma_e = 0.2$	$\sigma_e = 0.3$
N^{UI}	0.65	0.68	0.73
N^{SI}	0.59	0.60	0.62
$E(W^{UI} - W^{SI})$	4.1%	4.2%	4.3%
$\Pr(W^{UI} < W^{SI})$	0	0	0
$F_s(\Lambda)$	5%	14%	30%
$F_{u\varepsilon}(y)$	0.5%	1.2%	3.7%
spread	13%	14%	18%

In the above table, it is apparent that a more volatile economy results in higher employment along with a higher probability of default. These results are not surprising given the analysis earlier in the paper. In general in the model, volatility increases labour supply and output. Higher volatility also increases the credit spread and thus the benefit from its elimination.

5.1.3 Volatility of systemic risk

We now perform a similar experiment to the previous one only this time we drive up the volatility of the common, or *systemic*, financial shock, ε. As expected, the probability of a government bailout, $F_{u\varepsilon}(y)$, increases. However, our model predicts a negative relationship between the creadit spread and the volatility of the systemic shock. As a result, universal banking performs relatively less well when the volatility of the common shock increases.

Table 6. Systemic risk

$g = 0.2$	$\sigma_\varepsilon = 0.05$	$\sigma_\varepsilon = 0.1$	$\sigma_\varepsilon = 0.15$
N^{UI}	0.64	0.65	0.68
N^{SI}	0.58	0.59	0.60
$E(W^{UI} - W^{SI})$	4.2%	4.1%	4.0%
$\Pr(W^{UI} < W^{SI})$	0	0	0
$F_s(\Lambda)$	2.1%	5%	12%
$F_{u\varepsilon}(y)$	0.04%	0.5%	2.9%
spread	14.9%	13.3%	12%

5.1.4 Competition in investment banking sector

We now consider the case of relatively competitive investment banking. This is an interesting case to consider because it implies that the investment banking sector is more likely to be a source of volatility for the rest of the economy, other things constant, as the cushion of profits for absorbing potential losses is lower. On the other hand, increased competition reduces the double marginalization problem and therefore the relative gain from vertical integration. Increases in the parameter η are used to simulate the effects of increased competition. The table below shows that increased competition tends to reduce the average benefit of universal banking.

Table 7. Competition

$g = 0.2$	$\eta = 4$	$\eta = 6$	$\eta = 8$
N^{UI}	0.65	0.77	0.87
N^{SI}	0.59	0.68	0.74
$E(W^{UI} - W^{SI})$	4.1%	3.1%	2.0%
$\Pr(W^{UI} < W^{SI})$	0	0	5.1%
$F_s(\Lambda)$	5%	27%	53%
$F_{u\varepsilon}(y)$	0.5%	5.7%	17%

The intuition as to why increased competition might raise the riskiness of the investment banking sector seems straightforward: The lower mark-up shrinks the cushion of excess profits that absorbs the impact of low shock. That basic result seems, in spirit, consistent with arguments already in the literature that competition in financial markets may promote risk taking (see for example, Hellmann et al. (2000), Bolt and Tieman (2004), Repullo (2004) and Allen and Gale (2004)). The real issue, of course, is whether increased competition is welfare enhancing[11]. Thus higher risk may be positively correlated with higher investment, and therefore may promote production and welfare. That positive relation is also documented in some recent empirical work (see Claessence and Laeven, 2005). This is exactly what is observed in our model since employment, and therefore the output, is positively related to risk tolerance, reflected in the expected default threshold.

6 Conclusion

Should financial intermediaries and banks be broken up in the interests of financial and wider macroeconomic stability? Clearly, even in a simple model like the one just presented, the answer is far from straightforward as underlying distortions act to reinforce or offset other distortions. However our model outlines some key factors which are likely to

[11]For example, some microeconomic models suggest that competition not only increases risk but may also improve entrepreneurs' access to credit (Bolt and Tieman, 2004) and reduce the loan rate (Boyd and De Nicoló, 2005, Damjanovic, 2013).

be important (also in richer settings) in coming to an assessment of the desirability of separating deposit taking and risky investment activities.

Vertical integration implies higher production and lower prices for financial services, which in turn will result in higher consumption and welfare. Therefore the size of the economy is positively correlated with risk-taking. As noted in Kareken and Wallace (1978), that risk-taking may be excessive in the decentralised equilibrium. However, be that as it may, there is also a sense that risk-taking may be too low from an optimal policy perspective. That is because the decentralised economy is subject to monopolistic distortions and output is low relatively to the first-best. Encouraging banks to be more risky may actually be welfare-enhancing, ceteris paribus.

In our model, aligning banks' overall behaviour with the social good turns on a key trade-off; the eradication of a double marginalization problem (including a risk premium) in the financial sector, versus larger and costly government bail-outs. The bailouts per se may be welfare enhancing in the model as they ameliorate various underlying distortions. The degree of distortion of government action is important however. When such intervention is relatively efficient, our model suggests that universal banking may be preferable; double marginalization is more costly than government bail-outs. When government intervention is relatively distortive, then the interplay with shock variability may mean that separated banks are desirable. When the economy is competitive and the decentralized equilibrium features a relatively low degree of monopolistic distortion, vertical integration may be less desirable as in this case increases in risk-taking may be welfare decreasing.

References

[1] Allgood, Sam and Arthur Snow, 1998. The Marginal Cost of Raising Tax Revenue and Redistributing Income, Journal of Political Economy, 106 (6). pp. 1246-1273.

[2] Aksoy, Yunus and Henrique S. Basso and Javier Coto-Martinez, 2012. Lending Relationships and Monetary Policy, Economic Inquiry, 51 (1). pp. 368-393.

[3] Allen, Franklin and Douglas Gale, 2004. Competition and Financial Stability, Journal of Money, Credit and Banking, 36 (3).

[4] Angeloni, Ignazio and Faia, Ester, 2013. "Capital regulation and monetary policy with fragile banks," Journal of Monetary Economics, Elsevier, vol. 60(3), pages 311-324.

[5] Barth, Marvin J. III and Valerie A. Ramey, 2001. The cost channel of monetary transmission, NBER Macroeconomics Annual 16, 199–240.

[6] Benk, Szilard, Max Gillman and Michal Kejak. 2005. Credit Shocks in the Financial Deregulatory Era: Not the Usual Suspects, Review of Economic Dynamics, 8 (3) pp. 668-687.

[7] Benston, George J. 1994. Universal Banking, Journal of Economic Perspectives, 8 (3), pp. 121-144.

[8] Berger, Allen N., Gerald A. Hanweck, David B. Humphrey, 1987. Competitive viability in banking: Scale, scope, and product mix economies, Journal of Monetary Economics, 20 (3), pp. 501-520.

[9] Bernanke, Ben S., 1983. Nonmonetary effects of the financial crisis in the propagation of the Great Depression. The American Economic Review, 73 (3), pp. 257-276.

[10] Bernanke, Ben S., Mark Gertler, and Simon Gilchrist, 1999. The financial accelerator in a quantitative business cycle framework. In Handbook of Macroeconomics, vol. 1, edited by J. B. Taylor and M. Woodford, chap. 21. pp. 1341–1393.

[11] Boot, Arnould W. A and Anjan V. Thakor, Banking Scope and Financial Innovation, The Review of Financial Studies, 10 (4), pp. 1099-1131.

[12] Bolt, Wilko and Alexander F. Tieman, 2004. Banking Competition, Risk and Regulation, Scandinavian Journal of Economics, 106 (4), pp. 783-804.

[13] Boyd, John H and Chang, Chun and Smith, Bruce D, 1998. Moral Hazard under Commercial and Universal Banking, Journal of Money, Credit and Banking, Blackwell Publishing, 30 (3), pp. 426-68.

[14] Boyd, John H. and Gianni De Nicol, 2005. The Theory of Bank Risk Taking and Competition Revisited, The Journal of Finance, 60 (3), pp. 1329-1343.

[15] Brunnermeier, Markus K, Thomas M. Eisenbach and Yuliy Sannikov, 2012. Macroeconomics with Financial Frictions: A Survey. Princeton Working Paper.

[16] Carlstrom, Charles and Timothy Fuerst, 1997. Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis. American Economic Review 87 (5), pp. 893-910.

[17] Claessens, Stijn and Luc Laeven, 2005. Financial Dependence, Banking Sector Competition, and Economic Growth, Journal of the European Economic Association, 3 (1), pp. 179-207.

[18] Clark, Jeffrey, A. (1984). Estimation of Economies of Scale in Banking Using a Generalized Functional Form, Journal of Money, Credit, and Banking, 16 (1). pp. 53-68.

[19] Cornuejols, Gerard and Reha Tütüncü, 2007. Optimisation in Finance, Cambridge University Press.

[20] Christiano, Lawrence J., Roberto Motto, and Massimo Rostagno, 2014. Risk Shocks, American Economic Review, 104 (1), pp. 27-65.

[21] Diamond D.W., Dybvig P.H. ,1983. Bank runs, deposit insurance, and liquidity, Journal of Political Economy. 91 (3), pp. 401-419.

[22] De Fiore, Fiorella, Pedro Teles, and Oreste Tristani, 2011. Monetary Policy and the Financing of Firms, American Economic Journal: Macroeconomics, 3 (4), pp. 112-42.

[23] French, Kenneth R. et al., 2010. The Squam Lake Report: Fixing the Financial System. Princeton University Press.

[24] Gertler, M. and N. Kiyotaki, 2010. Financial Intermediation and Credit Policy in Business Cycle Analysis, Handbook of Monetary Economics, in: Benjamin M.

Friedman and Michael Woodford (ed.), Handbook of Monetary Economics, edition 1 (3), chapter 11, pp. 547-599.

[25] Gertler, M.; N. Kiyotaki and A. Prestipino, 2016. Wholesale Banking and Bank Runs in Macroeconomic Modelling of Financial Crises, NBER Working Papers 21892.

[26] Gertler, M.; N. Kiyotaki and A. Prestipino, 2017. A Macroeconomic Model with Financial Panics, Working paper.

[27] Goodfriend, Marvin and Bennett T. McCallum, 2007. Banking and interest rates in monetary policy analysis: A quantitative exploration. Journal of Monetary Economics, vol. 54 (5), pp. 1480-1507.

[28] Gorton, Gary. G., 2010. Questions and answers about the financial crisis, NBER Working paper 15787.

[29] Hancock, Diana, 1985. The Financial Firm: Production with Monetary and Non-monetary Goods, Journal of Political Economy, 93 (5), pp. 859-880.

[30] Hellmann, Thomas F., Kevin C. Murdock and Joseph E. Stiglitz, 2000. Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?," American Economic Review, vol. 90 (1), pp. 147-165.

[31] Hsieh, Chang-Tai and Peter J. Klenow, 2009. Misallocation and manufacturing TFP in China and India, Quarterly Journal of Economics, 124, pp. 1403-1447.

[32] Ireland, Peter N., 2014. The Macroeconomic Effects Of Interest On Reserves, Macroeconomic Dynamics, 18 (06), pp. 1271-1312, September.

[33] Kareken, John H. and Neil Wallace, 1978. Deposit insurance and bank regulation: A partial equilibrium exposition, Journal of Business, 51 (3), pp. 413-438.

[34] Kimball, Miles S., 1990. Precautionary Saving in the Small and in the Large, Econometrica, 58 (1), pp. 53-73.

[35] Loisel, O., 2014. Discussion of monetary and macroprudential policy in an estimated DSGE model of the euro area, International Journal of Central Banking, 10 (2), pages 169-236.

[36] Restuccia, Diego and Richard Rogerson, 2008. Policy Distortions and Aggregate Productivity with Heterogeneous Plants, Review of Economic Dynamics, 11, pp. 707–720.

[37] Repullo, Rafael, 2004. Capital requirements, market power, and risk-taking in banking, Journal of Financial Intermediation, . 13 (2), pp. 156-182.

[38] Spengler, J.J., 1950. Vertical Integration and Antitrust Policy, Journal of Political Economy, pp. 347-352.

[39] Thomas, Ewart A. S., 1971. Sufficient conditions for monotone hazard rate: An application to latency-probability curves, Journal of Mathematical Psychology, 8 (3), pp. 303–332.

7 Appendix

7.1 Existence, uniqueness and the second order conditions for the investment banking problem

In the text the following equation was derived to characterize optimal behaviour in the investment banking sector via a choice of Λ:

$$\int_{\Lambda}^{+\infty} \left[\frac{\eta - 1}{\eta} \frac{s}{\Lambda} - 1 \right] f^{s}(s)\, ds = 0. \tag{45}$$

First, it is established that no such solution for Λ need exist, and then that there may be many solutions. As an example for which no solution exists, consider Pareto distribution with pdf $f(s) = as^{-a-1}$ for $s \geq 1$. It follows that (45) becomes

$$\int_{\Lambda}^{+\infty} \left[\frac{\eta - 1}{\eta} \frac{as^{-a}}{\Lambda} - as^{-a-1} \right] ds = \Lambda^{-a} \left(\frac{\eta - 1}{\eta} \frac{a}{a-1} - 1 \right). \tag{46}$$

That expression is always positive if $a > \eta$, and a solution to (45) does not exist. On the other hand, if $a = \eta$, any Λ is a solution to (45).

7.2 Existence

To establish general conditions for existence and uniqueness of a solution, it is clear that one needs additional structure on the distribution function.

Definition 1 *We call the number A the supremum of the domain of the pdf f, if $\forall x$, $x < A$. It follows that $F(x) < 1$, and $\lim\limits_{x \to A} F(A) = 1$. For the lognormal distribution $A = +\infty$.*

Definition 2 *For any cdf $F(x)$ with positive domain, we define the "inverse log hazard function"*

$$h_{il}(x) = \frac{(1-F(x))}{xf(x)}. \tag{47}$$

We call it this since $h_{il}(x) = 1/h_l(x)$, where $h_l(x) = \frac{xf(x)}{(1-F(x))}$ is the hazard function for $y = \ln(x)$. Indeed, $F(X) = \Pr(x < X) = \Pr(y < \ln X) = F_y(\ln X)$. One may also compute the relation between the pdf x and pdf y.

$$f(X) = \frac{f_y(\ln X)}{X}. \tag{48}$$

It follows therefore that

$$h_l(x) = \frac{xf(x)}{(1-F(x))} = \frac{f_y(\ln X)}{1-F_y(\ln X)}.$$

To prove existence we will need the following assumption concerning the distribution:

Assumption A1:

$$\lim_{x \to A} h_{il}(x) = \frac{(1-F(x))}{xf(x)} = 0.$$

Proposition 6 *There exists a solution to (45) if the inverse log hazard rate converges to zero at the supremum of the domain,*

$$\int_{\Lambda}^{A} \left[\frac{s}{\Lambda} - \frac{\eta}{\eta-1}\right] f(s)\, ds = 0. \tag{49}$$

Proof. Consider the function

$$g_{30}(x) := \frac{\int_x^A \left[s - x\frac{\eta}{\eta-1}\right] f(s)\, ds}{x(1-F(x))} = \frac{\int_x^A sf(s)\, ds}{(1-F(x))x} - \frac{\eta}{\eta-1}.$$

It is easy to see that $\lim_{x \to 0} g_{30}(x) = \lim_{x \to 0} \frac{Es}{x} = +\infty > 0$. That implies that there exists a $c > 0$, such that for any $x \leq c$, $g_{30}(x) > 0$. We now wish to prove that

$$\lim_{x \to A} g_{30}(x) < 0.$$

Hence, we compute

$$\lim_{x \to A} g_{31}(x) = \frac{\int_x^A sf(s)\, ds}{(1-F(x))x}.$$

Since both the numerator and the denominator converge to 0 and are differentiable, one may establish if L'Hôpital's rule can be applied. Thus,

$$\lim_{x \to A} g_{31}(x) = \frac{\int_x^A sf(s)\, ds}{(1-F(x))x} = \frac{xf(x)}{xf(x) - (1-F(x))} = \frac{1}{1 - \frac{(1-F(x))}{xf(x)}}.$$

And if $\lim_{x\to A}\frac{(1-F(x))}{xf(x)}=0$ the limit exists and $\lim_{x\to A} g_{31}(x)=1$. Thus

$$\lim_{x\to A} g_{30}(x)=\lim_{x\to A} g_{31}(x)-\frac{\eta}{\eta-1}=-\frac{1}{\eta-1}. \tag{50}$$

By the definition of limit, one concludes that there exist an x^* such that for any $x\in[x^*, A)$, $g_{31}(x)<-\frac{0.5}{\eta-1}$. ■

Corollary 7 *If Assumption A1 is true, and x is the largest solution to $g_{30}(x)=0$ then $\forall$ $x_1>x$ we have that $g_{30}(x_1)\leq 0$.*

Proof. We will give a proof by contradiction. Assume that there is a solution, x, such that $g_{30}(x)=0$ and there exists $x_1>x$, such that $g_{30}(x_1)>0$. However, since Assumption A1 holds, formula (50) obtains, and there is a solution x_2 such that $g_{30}(x_2)=0$ and $x_2>x_1>x$. Therefore x is not the largest solution, and we have a contradiction. ■

From corollary 7 one also concludes that if x is the largest solution, $g'_{30}(x)\leq 0$ and function $g_{30}(x)$ cannot change sign from negative to positive at x.

7.2.1 Uniqueness and the second order conditions

Now, we may formulate a sufficient condition for uniqueness of the solution to $g_{30}(x)=0$.

Assumption A2: The inverse log hazard rate, $\frac{(1-F(x_2))}{x_2 f(x_2)}$, is a strictly decreasing function.

Corollary 8 *If distribution F satisfies Assumptions A1 and A2, then function $g_{30}(x)=0$ changes sign only once from positive to negative.*

Proof. We will give a proof by contradiction. Assume that there is a solution, x, such that $g_{30}(x)=0$; and that $g_{30}(x)=0$ changes sign only once from negative to positive. Then 1) $g'_{30}(x)\geqslant 0$ and therefore

$$g'_{30}(x)=\frac{-xf(x)\left[(1-F(x))x\right]-\left[(1-F(x))-xf(x)\right]\int_x^A sf(s)\,ds}{\left[(1-F(x))x\right]^2}\geqslant 0. \tag{51}$$

As x is a solution, we can rewrite (51)

$$g'_{30}(x)=\frac{xf(x)}{(1-F(x))x}\left(\frac{1}{\eta-1}-\frac{(1-F(x))}{xf(x)}\right)\geqslant 0.$$

From Corollary 7 we know that there exist $x_2>x$, such that $g_{30}(x_2)=0$, and $g'_{30}(x_2)\leq 0$. That implies that

$$\frac{1}{\eta-1}-\frac{(1-F(x_2))}{x_2 f(x_2)}\leqslant 0\leqslant\frac{1}{\eta-1}-\frac{(1-F(x))}{xf(x)};$$

or that

$$\frac{(1-F(x))}{xf(x)} \leqslant \frac{(1-F(x_2))}{x_2 f(x_2)},$$

which contradicts Assumption A2. ■

Function $g_{30}(x)$ is continuous and Corollary 8 implies that it changes sign only once from positive to negative. Assumption A2 implies that there can be only one solution with $g_{30}'(x) = 0$ therefore we can claim that there are no more than 2 solutions but only one of them corresponds to the situation when $g_{30}(x)$ changes sign from positive to negative. That solution will also satisfy the second order conditions of the initial problem.

Proposition 9 *If the distribution satisfies Assumptions A1 and A2, there is a unique solution x to $g_{30}(x) = 0$ at which function $g_{30}(x)$ changes sign from positive to negative. Only at this solution are both the first and the second order conditions satisfied.*

7.2.2 Lognormal distribution

It remains now to verify that the lognormal distribution satisfies assumptions A1 and A2. A1 asserts that

$$\lim_{x\to\infty} \frac{(1-F_y(x))}{f_y(x)} = 0. \tag{52}$$

for normal distribution. Applying L'Hôpital's rule, it follows that

$$\lim_{x\to\infty} \frac{(1-F_y(x))}{f_y(x)} = \lim_{x\to\infty} -\frac{f_y(x)}{f_y'(x)} = \lim_{x\to\infty} \left[-\frac{d}{dx}\ln(f_y(x))\right]^{-1}.$$

Hence, one needs to verify that for the normal distribution it is the case that

$$\lim_{x\to\infty} \left[-\frac{d}{dx}\ln(f_y(x))\right]^{-1} = 0. \tag{53}$$

And since $f_y(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$; and $\left[-\frac{d}{dx}\ln(f_y(x))\right]^{-1} = \sigma^2/(x-\mu)$, condition (53) is true for the lognormal function and A1 is satisfied.

Moreover, Thomas (1971) shows that the normal distribution has an increasing hazard rate. Therefore its inverse hazard rate is a decreasing function and assumption A2 is satisfied. It follows that for the lognormal distribution the solution exists. Moreover, the solution is unique and the second order conditions are satisfied.

7.3 Proof of Proposition 1

Proposition 1 $\Gamma^b(u\varepsilon)$ *is an increasing and concave function. Therefore more favourable systemic or macro shocks bring average returns closer to the contractual return. Moreover* $\Gamma^b(0) = 0$ *and* $\lim_{x\to\infty}\Gamma^b(x) = 1$.

$$\frac{d\Gamma^b(u\varepsilon)}{du\varepsilon} = \int_0^{\left(\frac{\Lambda}{u\varepsilon}\right)^{\frac{\eta}{\eta-1}}} \frac{e^{1-1/\eta}}{\Lambda} f^e(e)de > 0. \tag{54}$$

$$\frac{d^2\Gamma^b(u\varepsilon)}{d(u\varepsilon)^2} = -\frac{\eta}{\eta-1}\frac{1}{(u\varepsilon)^2} f^e\left[\left(\frac{\Lambda}{u\varepsilon}\right)^{\frac{\eta}{\eta-1}}\right]\left(\frac{\Lambda}{u\varepsilon}\right)^{\frac{\eta}{\eta-1}} < 0. \tag{55}$$

We apply L'Hôpital's rule to prove the last statement in the proposition.

7.4 Existence of commercial banks' default threshold

In this appendix we prove Proposition 2. First we show that there is a solution to (29)

$$g_{12}(z) := \frac{\int_z^{+\infty} \Gamma^b(x) f_{u\varepsilon}(x)dx}{(1 - F_{u\varepsilon}(z))} - \frac{\delta}{\delta - 1}\Gamma^b(z) = 0. \tag{56}$$

It is useful to establish some basic properties of the function $g_{12}(z)$. We do this in:

Lemma 10

$$\lim_{z\to\infty} g_{12}(z) = -\frac{1}{\delta - 1} < 0; \tag{57}$$

To prove the Lemma we apply L'Hôpital's rule:

$$\lim_{z\to\infty} \frac{\int_z^{+\infty} \Gamma^b(x) f_{u\varepsilon}(x)dx}{(1 - F_{ue}(z))} = \lim_{z\to\infty} \frac{\Gamma^b(z) f_{u\varepsilon}(z)}{f_{ue}(z)} = \lim_{z\to\infty} \Gamma^b(z) = 1.$$

Proposition 11 *There exists a solution to (29)*

It is easy to see that $g_{12}(0) \geq 0$. However from Lemma (10), there exists a z^* such that $\forall z > z^*$, $g_{12}(z) < -\frac{0.5}{\delta-1}$. Moreover, as $g_{12}(z)$ is a continuous function, there is a solution at which $g_{12}(y) = 0$ since $g_{12}(y)$ changes sign from positive to negative.

Corollary 12 *If y is the biggest solution to $g_{30}(y) = 0$ then $\forall\ x_1 > y$, $g_{12}(x_1) < 0$.*

We will give a proof by contradiction. Assume that there is a solution, y, such that $g_{12}(y) = 0$ and there exists $x_1 > y$, such that $g_{12}(x_1) > 0$. However, because of formula (57) there is a solution x_2 such that $g_{12}(x_2) = 0$ and $x_2 > x_1 > y$. Therefore y is not the largest solution.

7.5 The second order condition of the commercial bank problem

The initial problem is

$$\max_{R_t^c(i),y} E\Psi_{t+1} = \left[\int_y^{+\infty}\left[\frac{R_t^c(i)}{R_t^h}\Gamma^b(x)-1\right]f_{u\varepsilon}(x)dx\right]\left(\frac{R_t^c(i)}{R_t^c}\right)^{-\delta}B_t^C R_t^h; \tag{58}$$

$$\text{s.t.} : \frac{R_t^c(i)}{R_t^h}\Gamma^b(y)=1.$$

To proceed, we substitute in the constraint so that

$$\max_y E\Psi_{t+1}(y) = \left[\int_y^{+\infty}\left[\frac{\Gamma^b(x)}{\Gamma^b(y)}-1\right]f_{u\varepsilon}(x)dx\right]\left(\Gamma^b(y)\right)^{\delta}B_t^C R_t^h\left(\frac{R_t^h}{R_t^c}\right)^{-\delta}. \tag{59}$$

The first order condition is

$$\frac{d}{dy}E\Psi_{t+1}(y) = \left[\int_y^{+\infty}\left((\delta-1)\Gamma_{u\varepsilon}(x)-\delta\Gamma_{u\varepsilon}(y)\right)f_{u\varepsilon}(x)dx\right]\left(\Gamma^{b\prime}(y)\right)\left(\Gamma^b(y)\right)^{\delta-2}B_t^C R_t^h\left(\frac{R_t^h}{R_t^c}\right)^{-\delta}=0. \tag{60}$$

or

$$\frac{d}{dy}E\Psi_{t+1}(y) = g_{12}(y)((\delta-1)\left(\Gamma^{b\prime}(y)\right)\left(\Gamma^b(y)\right)^{\delta-2}B_t^C R_t^h\left(\frac{R_t^h}{R_t^c}\right)^{-\delta}(1-F_{ue}(y)).$$

and has the same sign as $g_{12}(y)$. The second order conditions are satisfied if and only if $g_{12}(y)$ changes sign from positive to negative. However, we have proved that such a y always exists.

7.6 Closed form solution

Here we show the derivation of the closed form solution to model presented in Table 2.

First, begin by rewriting some of the equations from Table 2 in the main text as follows,

$$C_t = \frac{1}{\lambda}\frac{1}{\mu^F}\Omega_t A_t^\rho \frac{\Lambda}{R_t^C}(\Delta)^{1/\eta}; \tag{61}$$

$$Y_{t+1} = A_t^\rho u_{t+1}\Omega_t\Delta\varepsilon_{t+1}N_t; \tag{62}$$

$$G_{t+1}^{SI} = N_t W_t R_t^h \Gamma^{gJ}(u_{t+1}\varepsilon_{t+1}) = N_t\frac{1}{\mu^F}\Omega A_t^\rho\Lambda(\Delta)^{1/\eta}\frac{\Gamma_{u\varepsilon}^{gJ}(u_{t+1}\varepsilon_{t+1});}{\text{sp}^J} \tag{63}$$

$$C_{t+1} = Y_{t+1}-gG_{t+1}; \tag{64}$$

$$R_t^C = \text{sp}^J R_t^h. \tag{65}$$

Combining (61) with (65) results in (66) and substitution of (62) and (63) into (64) gives (67)

$$R_t^h C_t = \frac{1}{\lambda}\frac{1}{\mu^F}\Omega_t A_t^\rho\Lambda(\Delta)^{1/\eta}\frac{1}{\text{sp}^J}; \tag{66}$$

$$C_{t+1} = A_t^\rho N_t\Omega_t\Delta\left[u_{t+1}\varepsilon_{t+1}-\widetilde{g}\frac{\Gamma^{gJ}(u_{t+1}\varepsilon_{t+1});}{\text{sp}^J}\right]. \tag{67}$$

Where $\widetilde{g} := g\frac{1}{\mu^F}(\Delta)^{1/\eta-1}\Lambda$. Combine the previous two equations with the consumption Euler equation (M1) to obtain (37).

7.7 Proof of Proposition 5 about employment under insurance

Here we consider insured separated banking versus insured universal banking. This is an interesting comparison, with some real-world resonance. It will turn out that labour supply is unambiguously higher under universal banking, that is regardless of g and $u\varepsilon$. As to whether universal banking is still *preferable* to separated banking, that will depend on whether $N^{UI} \leq N^{*}$. To show that $N^{UI} > N^{SI}$, one needs to prove that $\Upsilon^{UI} > \Upsilon^{SI}$, where

$$\Upsilon^{UI} = \int_{0}^{+\infty} \frac{\Gamma^{g}(u\varepsilon) + \Gamma^{b}(u\varepsilon)}{u\varepsilon - \widetilde{g}\Gamma^{g}(u\varepsilon)} dF_{u\varepsilon}$$

$$\Upsilon^{SI} = \int_{0}^{+\infty} \frac{\Gamma^{gS}(u\varepsilon) + \Gamma^{bS}(u\varepsilon)}{u\varepsilon/\Gamma^{b}(y) - \widetilde{g}\Gamma^{gS}(u\varepsilon)} dF_{u\varepsilon}.$$

As noted, for any realisation of the common shock we can prove that

$$\frac{\Gamma^{g}(u\varepsilon) + \Gamma^{b}(u\varepsilon)}{u\varepsilon - \widetilde{g}\Gamma^{g}(u\varepsilon)} > \frac{\Gamma^{b}(y)\left[\Gamma^{gS}(u\varepsilon) + \Gamma^{bS}(u\varepsilon)\right]}{u\varepsilon - \widetilde{g}\Gamma^{gS}(u\varepsilon)\Gamma^{b}(y)}.$$

To prove that, we first show $\Gamma^{g}(u\varepsilon) + \Gamma^{b}(u\varepsilon) > \Gamma^{b}(y)\left[\Gamma^{gS}(u\varepsilon) + \Gamma^{bS}(u\varepsilon)\right]$ in Lemma 13. Then in lemma 14, which is a statement about the size of bailouts across different banking structures, we prove that $\Gamma^{g}(u\varepsilon) > \Gamma^{gS}(u\varepsilon)\Gamma^{b}(y)$. That establishes that for any common shock $u\varepsilon$ and for any value of the government inefficiency, g,and therefore that $\Upsilon^{UI} > \Upsilon^{SI}$.

Lemma 13 *For any realisation $u\varepsilon$,*

$$\Gamma^{g}(u\varepsilon) + \Gamma^{b}(u\varepsilon) > \Gamma^{b}(y)\left[\Gamma^{gS}(u\varepsilon) + \Gamma^{bS}(u\varepsilon)\right]. \tag{68}$$

Proof. First recall the definitions $\Gamma^{g}(u\varepsilon) = \min\left(\frac{g}{\widetilde{g}}\frac{u\varepsilon}{1+g}; \left(1 - \Gamma^{b}(u\varepsilon)\right)\right)$. Therefore

$$\Gamma^{g}(u\varepsilon) + \Gamma^{b}(u\varepsilon) = \min\left(\frac{g}{\widetilde{g}}\frac{u\varepsilon}{1+g} + \Gamma^{b}(u\varepsilon); 1\right). \tag{69}$$

The other definitions are

$$\Gamma^{gS}(u\varepsilon) + \Gamma^{bS}(u\varepsilon) = \min\left(\frac{g}{\widetilde{g}}\frac{u\varepsilon}{\Gamma^{b}(y)}\frac{1}{1+g} + \Gamma^{bS}(u\varepsilon); 1\right), \tag{70}$$

where $\Gamma^{bS} = \min\left(\frac{\Gamma^{b}(u\varepsilon)}{\Gamma^{b}(y)}, 1\right)$. Multiply (70) by $\Gamma^{b}(y)$ so that

$$\Gamma^{b}(y)\left[\Gamma^{gS}(u\varepsilon) + \Gamma^{bS}(u\varepsilon)\right] = \min\left(\frac{g}{\widetilde{g}}\frac{u\varepsilon}{1+g} + \min(\Gamma^{b}(u\varepsilon), \Gamma^{b}(y)); \Gamma^{b}(y)\right). \tag{71}$$

Now we need to compare (71) with (69). Indeed

$$\min\left(\frac{g}{\widetilde{g}}\frac{u\varepsilon}{1+g} + \Gamma^{b}(u\varepsilon); 1\right) > \min\left(\frac{g}{\widetilde{g}}\frac{u\varepsilon}{1+g} + \min(\Gamma^{b}(u\varepsilon), \Gamma^{b}(y)); \Gamma^{b}(y)\right)$$

as $\frac{g}{\widetilde{g}}\frac{u\varepsilon}{1+g} + \Gamma^{b}(u\varepsilon) > \frac{g}{\widetilde{g}}\frac{u\varepsilon}{1+g} + \min(\Gamma^{b}(u\varepsilon), \Gamma^{b}(y))$, and $1 > \Gamma^{b}(y)$ ■

Lemma 14 *For any realisation $u\varepsilon$,*

$$\Gamma^{g}(u\varepsilon) > \Gamma^{gS}(u\varepsilon)\Gamma^{b}(y). \tag{72}$$

Proof. We recall the definitions

$$\Gamma^{g}(u\varepsilon) = \min\left(\frac{g}{\widetilde{g}}\frac{u\varepsilon}{1+g}; \left(1 - \Gamma^{b}(u\varepsilon)\right)\right);$$

$$\Gamma^{b}(y)\Gamma^{gS}(u\varepsilon) = \min\left(\frac{g}{\widetilde{g}}\frac{u\varepsilon}{1+g}; \left(\Gamma^{b}(y) - \min\left[\Gamma^{b}(u\varepsilon); (\Gamma^{b}(y)]\right)\right);$$

to complete the proof we just need to show that $1 - \Gamma^{b}(u\varepsilon) > \Gamma^{b}(y) - \min\left[\Gamma^{b}(u\varepsilon); (\Gamma^{b}(y)\right]$; which is certainly correct for both cases: when $y < u\varepsilon$ and when $y > u\varepsilon$. ■

It follows immediately from Lemmas 13 and 14 that $\Upsilon^{UI} > \Upsilon^{SI}$, and therefore that $N^{UI} > N^{SI}$, which completes the proof of Proposition 5.