Bank size, risk diversification and money markets*

Hugo Rodríguez Mendizábal†
Instituto de Análisis Económico (CSIC) and Barcelona GSE

April 21st, 2016
Very preliminary and incomplete.
Please, do not distribute or cite without permission.

Abstract

This paper presents a theoretical model based on risk diversification to rationalize the observed dichotomy in money markets by which small banks are net providers of funds while large banks become net purchasers. Unlike the existing literature on this topic, the model incorporates liquidity provision by a central bank. In the model, smaller banks are less diversified and more risky which means producing a lower amount of loans through smaller leverage and borrowing in the wholesale money market with larger risk premiums. Because payment needs for settlement purposes are random and because smaller banks face worse rates in the interbank market, in equilibrium they will obtain from the central bank extra funds for precautionary reasons and offer these excess reserves in the money market. The opposite will be true for large banks.

Keywords: bank size, diversification, money market, bank solvency
JEL classification codes: E4, E5, G21

1 Introduction

This paper presents a model of the money market dichotomy between large and small banks. The empirical literature has repeatedly found that small banks tend to be net sellers of funds while large banks tend to be net buyers

*I would like to thank Jordi Caballe, Joachim Jungherr, Gabriel Pérez Quirós and Jaume Ventura for helpful dicussions. I would like to acknowledge the support of the Barcelona GSE Research Network as well as the financial support of the Spanish Ministry of Science and Competition through grant ECO2012-38460 and of the Generalitat de Catalunya through Grant 2014SGR-1446. I am also grateful for financial support from the ADEMU project, "A Dynamic Economic and Monetary Union", funded by the European Union's Horizon 2020 Program under grant agreement No 649396. All errors are my responsibility.

†Instituto de Análisis Económico. Campus UAB. 08193 Bellaterra Barcelona, Spain. E-mail: hugo.rodriguez@iae.csic.es.

in the money market.[1] Allen et al. [2] review three possible reasons for this dichotomy. First, if small banks are more risk averse than large banks, as in Ho and Saunders [23], they would rely on deposits to finance their assets instead of on the interbank market. Second, smaller banks may prefer deposit financing if they have lower deposit-taking costs than larger banks. As a third explanation, Allen and Saunders [3] show that adverse selection resulting from information asymmetries can also produce this observed pattern of trade if small banks are perceived as more opaque or riskier than large banks.[2]

The purpose of this paper is to show that the small-large dichotomy in the money market can be reproduced by a theoretical model based on diversification opportunities associated with size. Unlike the previous literature, in my model banks are risk neutral institutions and there are no information asymmetries regarding default probabilities or differences in deposit-taking costs. In sum, banks are treated symmetrically apart from the assumption that larger banks are more diversified and face lower risks than smaller banks. Furthermore, and also unlike the previous literature, my model economy incorporates liquidity provision by a central bank. Thus, commercial banks may decide to obtain the liquidity they need first hand from the central bank through an open market operation, or borrow it second hand, on an unsecured basis, from another commercial bank through the interbank market.

In the model, because of the existence of agency costs, being less risky allows large banks to produce relatively more loans through larger leverage. Furthermore, they will have to pay a smaller risk premium when borrowing on an unsecured basis in the interbank market. In contrast, smaller banks will be less diversified and more risky which means producing a lower amount of loans and borrowing in the wholesale money market with larger risk premiums. As it will be clear below, producing more loans also implies a larger demand of funds for payment settlement purposes. Because these payment needs are random and because smaller banks face worse rates in the interbank market, in equilibrium they will demand extra funds for precautionary reasons in the open market operation and offer these excess reserves in the interbank market. The opposite will be true for large banks. Thus, larger banks will be net buyers of funds in money markets and smaller banks net sellers.

In building up the model, this paper makes a stand about how money is created in our economies and the role commercial banks, monetary authorities, and interbank markets play in this process. In the traditional view to which the papers mentioned above belong, banks act as intermediaries between depositors and borrowers. The bank lending process starts when a saver deposits an asset on a bank. After the deposit is made, the bank then loans out those assets to borrowers. Should a bank need more funds to lend to additional borrowers,

[1] See, among others, Gambs and Kimball [20], Furfine [18], Allen et al. [2], and, more recently, Cocco et al. [13], Craig and von Peter [15] or the anecdotal evidence described in Stigum and Crescenzi [15].

[2] There are other theoretical papers analyzing this dichotomy. Chen and Mazumdar [11] combine the three explanations described by Allen et al. [2]. Ashcraft et al. [4] impose this dichotomy by assuming credit constraints and limited participation for small banks.

it can search for them in the interbank market. In this case, banks that are not using all their deposits for loans to customers, can provide them to needed banks. The existing explanations of the money market dichotomy then rely on the perceived costs and risks of these different forms of funding across banks with different sizes.

As an aside, the money multiplier is one particular way in which this traditional view considers bank intermediation, with the aim of linking broad (i.e. M1) with narrow (i.e. M0) definitions of money. According to the money multiplier, money creation starts when a client deposits outside money in a bank account. Because banks have to satisfy reserve requirements, a fraction of the deposits they maintain has to be in the form of reserves but the reminder can be lent out. These loans are then deposited by other agents at different credit institutions which themselves lend out a fraction of those deposits. Iterating this argument forward, a given initial amount of outside money multiplies itself into a much larger amount of bank lending and deposits.

Three important results arise from both the general traditional view on bank intermediation as well as its particular money multiplier interpretation: (i) the level of deposits (or the level of reserves) in the banking system acts as a constraint on the amount of intermediation to be done by banks, (ii) the amount of deposits in banks is seen as the origin of the amount of credit banks can produce, and (iii) interbank money markets as well as markets for loans to nonfinancial institutions and households all deal with the same homogeneous object, namely, the assets that initially are deposited by savers in commercial banks.

This traditional view of bank intermediation seems at odds with several institutional conventions associated with banking in modern economies. First, as a matter of practice, commercial banks create money, in the form of bank deposits, when making new loans. When a bank makes a loan, it credits the client's bank account with a bank deposit of the size of the loan. Thus, the bank does not "search" for a depositor from whom to transfer funds to a borrower. It just creates the new deposit out of the thin air. This view is shared both by "orthodox" academicians (see Goodhart [22]), as well as central bankers (see, among others McLeay et al. [27], from the Bank of England, Holmes [26], from the Federal Reserve Bank of New York, or Constancio [14] from the ECB) and market practitioners (see Sheard [28]). Importantly, adopting this alternative mechanism, dubbed "fountain pen" by Tobin [30], implies that it is loans, and the incentives to create credit, what originates deposits and not the other way around, as the traditional view contends.

From an accounting perspective, loan creation in reality produces a bank asset on the amount of the loan together with a bank liability associated with the deposits, in terms of nominal units of account, the loan creates. Notice that, initially when the loan is produced, both the loan and the deposit refer to the same agent. Then, when this agent wants to dispose of its deposits to make a payment (which is the fundamental reason the agent asked for a loan in the first place) two things may happen. If the recipient of the payment has an account in the same bank as the payer, the bank just rename those deposits. Notice now the asset and liability in the bank balance sheet refer to two different agents but the timing is reversed as compared with the traditional view. On the other hand, if the recipient of the payment is not a client of the same bank as the payer, funds need to be transferred to another depository institution. These funds are usually reserves, that is, current accounts of depository institutions at the central bank. In this case, two new double entries are recorded by which both deposits and reserves are reduced at the credited bank and increased at the recipient bank.[3]

This discussion leads to a second challenge of the traditional view, in particular, of its money multiplier interpretation: the role reserves play in the process of money creation. Within the multiplier view, reserves are just a mechanical way to connect the monetary base to broad definitions of money through reserve requirements. In this sense, this view is silent regarding its application to countries with no reserve requirements. On the contrary, in reality banks maintain reserves for two reasons. One is to satisfy reserve requirements wherever these requirements are in place. However, these requirements are typically defined over previously held deposits and, therefore, are interest rate insensitive. The second reason is to satisfy depositor's payments demand. Whenever a client wishes to make a payment to be done at another bank, this payment is usually done with reserves. Thus, reserve demand is driven by the netting of payments derived from the loan and deposit creation to finance economic activity. Notice this "netting of payments" should also include deposit competition by which one bank tries to attract other banks' deposits.

In general, central banks define their monetary policy stance as a target on very short rates (i.e. overnight) in money markets. In implementing its monetary policy, monetary authorities are usually ready to supply, at the target rate, as much reserves as depository institutions demand. When the monetary policy stance changes, the central bank modifies its interest rate target but still "reads" the amount of reserves needed to support that new target from demand by commercial banks. The arguments in this and the previous paragraphs then suggest that both deposits and reserves are a consequence of the loan production and, therefore, cannot constrain loan creation by banks as is implicit in the traditional view.

The role of reserves in our banking systems also generates a third observation regarding the traditional view. In reality, reserves, in the form of current accounts at the central bank, are a completely different object than customer's deposits at commercial banks or the loans these banks provide to their customers. Reserves are produced by the central bank, while loans and deposits are produced by commercial banks. As we have seen before, reserves are a consequence of loan and deposit creation by banks. Thus, it does not make any sense to claim that depository institutions can use the interbank market to transfer among them the assets deposited by their clients to fund loans to the nonfinancial sector. Furthermore, this argument also questions the widespread

[3] The accounting conventions that allow banks to create money out of nothing are described in Werner [32] and [33].

view by which banks can "lend out" their reserves to nonfinancial agents. As an example, part of the current debate about QE has to do with the inflationary pressures that would occur if financial institutions were to lend out to nonfinancial firms and households the massive amounts of excess reserves stored at central banks. This claim is inaccurate as for that to be the case depository institutions would have to transfer those reserves to nonfinancial institutions and households. This is not possible as these agents typically do not hold accounts at the central bank (with possibly the exception of the Treasury). Individual banks could dispose of their reserves by lending them to another depository institution through the interbank market or by using them to buy an asset from another bank. But in the aggregate, at any particular point in time, the amount of reserves held by depository institutions is fixed as these transfers only reshuffle them between individual banks. They cannot move out to the nonfinancial sector.[4]

From the discussion above, if official reserves and existing deposits do not constrain bank lending and money creation, what does? Banks face several constraints, some exogenous to the banking system and some endogenous. The most obvious exogenous constraint is capital requirements. Banks have to maintain a fraction of their risky assets in the form of capital. As capital is expensive to collect, the potential to create new loans (and the money associated with them) is impaired by the amount of capital banks hold. Another limit is loan demand. To create new money, someone has to agree to take a loan. Furthermore, banks themselves constrain their lending behavior as they look for profitable opportunities, at a reasonable level of risk, where to place their loans. Finally, all these constraints are also affected by monetary policy as it influences the opportunity cost of money together with the amount of liquidity with which to fund these loans and how the interbank market distributes it among depository institutions. All these dimensions in which endogenous money is created will be present in the model below.

This endogenous view of the process of money creation is slowly entering mainstream economics. For example, Disyatat [16] uses this approach when revisiting the bank lending channel of monetary policy. His model, however, is partial equilibrium. Jakab and Kumhof [24] compare the response to various shocks under the traditional view and the endogenous money creation view (what they call "financing through money creation") and conclude the way we model bank intermediation may have significant quantitative effects. Chari and Phelan [10] and Goodfriend and McCallum [21] also include some form of endogenous deposit creation by banks. These models, however, assume binding reserve requirements as a constraint on the deposit production by banks. As explained above, under the institutional arrangements we observe in reality, this does not seem to be the case. Compared with this recent literature, my model incorporates an explicit connection between monetary policy instruments

[4] Apart from the Treasury, the other exeption is cash holdings by the nonfinancial sector. As we withdraw cash from ATMs, banks use their reserves to get the banknotes needed to replenish them. However, if these movements are significant, the central bank usually intervenes to restore the levels of reserves held previously by banks.

(as summarized by open market operations, standing facilities and reserve requirements), bidding behavior by banks at the open market operations, and net positions in the interbank market and links these items to the funding of decisions on real economic activity such as employment and production.

The remaining of the paper is organized as follows. Section 2 presents the model. Section 3 characterizes the equilibrium. Finally, Section 4 concludes.

2 The Theoretical Model

2.1 The setup

The economy consists of a continuum of identical islands with measure 1. Figure 1 depicts one such islands represented by a circle with circumference length equal to 1. On each point of the circle there is a location which includes a measure 1 of risk averse workers and a measure 1 of risk neutral entrepreneurs. On every island there also exist two banks, one being large and the other small. All banks in the economy are owned by a measure 1 of risk neutral investors. Finally, there is a public sector composed of a central bank and a deposit insurance scheme which encompasses all islands.

There are two periods $t = 0, 1$. On period $t = 0$, agents make choices. On period $t = 1$ uncertainty is resolved. Workers and entrepreneurs start period $t = 0$ with no resources. Each worker inelastically supplies a unit of labor and entrepreneurs hire these workers. Capital is not needed for production. As it will be clear below, workers could also work as auditors, monitoring commercial banks in case these banks go bankrupt. The labor market is perfectly competitive within each island, with workers being able to work for entrepreneurs in any location.

Commercial banks connect entrepreneurs and workers. The size of each bank refers to the number of locations it serves. In this sense, the small bank is present in a measure δ_S of *contiguous* locations where $0 < \delta_S < 0.5$. The large bank serves the remaining $\delta_L = 1 - \delta_S$ locations where $0.5 < \delta_L = 1 - \delta_S < 1$. Thus, I am assuming that each location is served by just one bank and that agents can only have business with the bank serving their location. This means entrepreneurs can only ask for a loan to and the workers maintain deposits in the bank at the location they live. Without loss of generality, I take the large bank to serve, moving clockwise, from location 0 to location δ_L while the small bank serves the remaining $1 - \delta_L$, as it is shown in Figure 1.[5] As explained in the Introduction, unlike most of the previous Macroeconomic literature, these banks create means of payments (deposits) in the process of providing loans to entrepreneurs. Commercial banks are subject to solvency and liquidity risks. To face these risks, at $t = 0$ they collect capital provided by investors and use the liquidity facilities of the central bank as well as an economywide interbank market. Investors are not allocated to any specific location.

[5] Here the idea of location should not necessarily be taken literaly as a geographical place but as market segmentation along different bank sizes.

Finally, the central bank sets the opportunity cost of funds and solve liquidity problems of commercial banks. The deposit insurance scheme insures depositors against the solvency risk of commercial banks. The model has two building blocks: a real side and a nominal side. For easiness of exposition I will describe below each of these blocks separately.

2.2 Timing and balance sheets

To see the whole flow of funds, take a general bank with size δ, meaning serving a measure δ of contiguous locations. At the very beginning of period $t = 0$ this bank has no resources. Then, first, investors provide outside nominal assets (such as government bonds or foreign IOUs) which can be in the form of equity or deposits. For the time being, assume investors only provide equity to banks (this is what will happen in equilibrium). Let $Q(\delta)$ be the amount of equity *per location* this bank receives from investors. Thus, the initial balance sheet of this bank will read

Balance sheet of bank δ when portfolio decisions are taken by investors at $t = 0$

Assets			Liabilities
Assets	$\delta Q(\delta)$	Equity	$\delta Q(\delta)$

Once this investment is made, and still at period $t = 0$, the bank seeks liquidity at an open market operation (OMO) at the central bank. Let $M(\delta)$ be the allotment of such an operation which uses part of the initial bank's assets as collateral. Thus, the balance sheet will be

Balance sheet of bank δ when the open market operation takes place at $t = 0$

Assets			Liabilities
Reserves	$M(\delta)$	Equity	$\delta Q(\delta)$
Assets	$\delta Q(\delta) - M(\delta)$		

After the OMO, the bank provides loans to entrepreneurs of $L(\delta)$ on each of the locations it serves. When providing these loans, the bank makes a double entry in its books. On the asset side, the bank annotates the right associated with the loan taken by the entrepreneur. On the other hand, deposits (i.e., means of payments) are created, and the liability side reflects the right of entrepreneurs (obligation for the bank) to dispose of those deposits to make payments. Thus, the balance sheet of the bank at the time loans are granted on period $t = 0$ is:

Balance sheet of bank δ when loans are made at $t = 0$

Assets			Liabilities
Reserves	$M(\delta)$	Deposits (ent.)	$\delta L(\delta)$
Assets	$\delta Q(\delta) - M(\delta)$	Equity	$\delta Q(\delta)$
Loans (ent.)	$\delta L(\delta)$		

These loans are used by managers to pay workers for wages in advance. This means entrepreneurs order transferring the property of these deposits to workers. In this process there is payment uncertainty, though. Once it is resolved, banks know whether they have an excess or deficit of reserves and access the interbank market to compensate it. Let $I(\delta, h)$ be the interbank lending of the bank (borrowing if negative), which will depend upon the realization of a random variable h to be specified below. The balance sheet now reads

Balance sheet of bank δ when interbank market opens at $t = 0$

Assets			Liabilities
Reserves	$M(\delta) - I(\delta, h)$	Deposits (ent.)	$\delta L(\delta)$
Assets	$\delta Q(\delta) - M(\delta)$	Equity	$\delta Q(\delta)$
Loans (interbank)	$I(\delta, h)$		
Loans (ent.)	$\delta L(\delta)$		

Finally, the transfers are made and workers keep these wages as deposits in the bank until the following period. The period $t = 0$ ends after these transfers are made. Thus, the balance sheet of the bank at the end of period $t = 0$ is:

Balance sheet of bank δ when interbank market opens at $t = 0$

Assets			Liabilities
Reserves	$M(\delta) - I(\delta, h)$	Deposits (work.)	$\delta L(\delta)$
Assets	$\delta Q(\delta) - M(\delta)$	Equity	$\delta Q(\delta)$
Loans (interbank)	$I(\delta, h)$		
Loans (ent.)	$\delta L(\delta)$		

At the beginning of period $t = 1$, workers go to work for entrepreneurs and production takes place. It is at this time that production uncertainty is resolved which makes a random fraction of entrepreneurs in each location to produce nothing. Unlucky entrepreneurs default on their loans. Depending on the locations each bank serves, these defaults may make some banks go bankrupt. Auditors audit bankrupt banks while solvent banks pay returns on deposits and equity. Once all agents receive incomes, consumption takes place. Because there is no aggregate uncertainty, the fraction of insolvent banks at $t = 1$ is known at $t = 0$. Thus, at $t = 0$, agents know the amount of workers to be hired as auditors at $t = 1$ and, therefore, who will not be producing goods. These auditors will be paid the competitive wage.

Notice $\delta L(\delta)$ appears both in the asset and liability sides of the balance sheet not because the bank is intermediating between depositors (workers) and borrowers (managers). In fact, when the loans were created, managers appeared on both sides of the bank's balance. It is the loan creation to make payments what causes different agents to be on both sides of the bank's balance sheet. It is in this sense that loans create deposits and not the other way around.[6] Also

[6] As it will be clear below, all locations start identical at the time banks make their decisions. This means banks create the same amount of loans and deposits in each location they serve. Thus, the total balance sheet of a bank of size δ is just δ times the volume of the balance for each location.

notice interbank loans and loans to entrepreneurs are very different objects. Interbank loans are done with reserves, which are supplied by the central bank. Thus, an interbank loan does not change the size of the balance sheet of the bank. It just increases an asset category (loans in the interbank market) at the expense of the reduction in another category (reserves). In contrast, loans to entrepreneurs are produced by creating deposits and, therefore, change the size of the balance sheet by increasing both an asset (loans to entrepreneurs) and a liability category (deposits). Finally, notice reserve holdings are used to face liquidity risks from payments while equity is used to deal with solvency risks. These two risks are related through the total amount of loans $\delta L(\delta)$ which is a decision of banks. In solving these problems, the bank will first decide on the amount of loans, $L(\delta)$, considering solvency risks, and then will work out its implications for liquidity risks and reserve demand. Because both its reserve demand as well as the loans provided to entrepreneurs (and, therefore, the amounts of transfers to be made to workers) will endogenously vary with the size of each bank, δ, some banks will be net recipients of funds and will make them available in the interbank market while other banks will face net outflows of funds and will demand them in the interbank market. Below I will show how the interbank position of banks relates to bank size δ and what are the factors affecting it.

2.3 The nominal side (i): solvency risks

At $t = 0$ entrepreneurs on all islands hire workers on their islands with production taking place at $t = 1$. To incorporate a role for a banking sector, suppose direct trade between entrepreneurs and workers is not possible in the exchange of labor and consumption goods. This impossibility can be rationalized by assuming that, by the time workers are hired by entrepreneurs at $t = 0$, there is no commitment on the part of entrepreneurs to pay wages to workers after production takes place at $t = 1$. This friction can also occur if, as in Kiyotaki and Moore [25], commitment or resalability problems prevent the IOUs signed at $t = 0$ by entrepreneurs when hiring workers to be used by these workers at $t = 1$ to buy goods from other entrepreneurs. This may well happen if different locations produce different goods demanded by workers. In these scenarios there is demand for an institution that either has the technology to enforce contracts or that produces IOUs recognizable by all agents in the economy. This is the role of banks in this model when providing loans and creating deposits.

The particular way in which information or commitment problems are resolved by banks is as follows. Whenever an entrepreneur wants to hire workers at $t = 0$, he will ask for a loan to his local bank equal to his nominal wage bill. Then the entrepreneur will order the bank to transfer these funds to the accounts of the workers. For the time being, assume these workers' accounts are maintained in the same bank. Thus, the bank will only need to rename the owner of those deposits. This solves the commitment problem of the entrepreneurs as wages are paid in advance. On the other hand, it also solves any commitment problem on the part of workers as the receipt from these transfers is proof of the wage payments and, therefore, can be used by entrepreneurs to claim the workers' labor services. Additionally, deposits are homogeneous units of account that can be used by depositors to buy goods from any producer. When a depositor (worker) pays for consumption goods at $t = 1$, he transfers these deposits to an entrepreneur. Entrepreneurs then use these funds at $t = 1$ to pay back the loan they asked for at $t = 0$. Within this process, banks face solvency risks as entrepreneurs may fail to pay back their loans at $t = 1$. Banks confront this solvency risk by accumulating capital provided by investors at the beginning of $t = 0$.

2.3.1 Investors

As mentioned above, in the economy there is a measure 1 of identical, risk neutral investors. They start period $t = 0$ with nominal assets A, which are distributed between bank equity, Q, earning a gross nominal return $1 + \iota^q$, bank deposits, D, with gross nominal return $1 + \iota^d$, or central bank certificates of deposits, O, paying interest i^o, i.e.

$$Q + D + O \leq A.$$

In general, total equity and deposits are distributed across all banks in the economy. Being risk neutral, investors make these portfolio decisions to maximize their net worth at $t = 1$, A', that is,

$$A' = (1 + \iota^q)Q + (1 + \iota^d)D + (1 + i^o)O.$$

Because investors diversify across all banks and islands, the total return on bank equity and deposits is deterministic even though some of the banks are insolvent along the way.

2.3.2 Banks

As all banks start identical, we can look at the problem of a general bank at a representative island. Take one of these islands and a bank of size $\delta \in [0, 1]$. Let $x(\delta) \in [0, 1]$, be the fraction of entrepreneurs served by this bank who are able to pay back the bank loan. Since all entrepreneurs served by this very bank are exante identical, they will all ask for the same loan amount and the NPL ratio of that bank will be $1 - x(\delta)$. Assume this fraction $x(\delta)$ is a random variable with marginal distribution denoted $\Xi(x; \delta)$ and corresponding density $\xi(x; \delta)$. Let $\mu(\delta)$ be the mean of $\Xi(x; \delta)$. I impose the following assumptions.

Assumption 1 $\Xi(x; \delta)$ is such that:

(i) $\Xi(x; \delta)$ is continuously differentiable with respect to both x and δ, for all $x \in [0, 1]$ and $\delta \in [0, 1]$,

(ii) $\mu(\delta) = \mu$, for all $\delta \in [0, 1]$,

(iii) for any δ_S and δ_L such that $\delta_S < \delta_L$, and any $a \in [0,1]$

$$\int_0^a \Xi(x;\delta_L)dx \leq \int_0^a \Xi(x;\delta_S)dx$$

with strict inequality over some interval in $[0,1]$,

(iv) the generalized hazard rate

$$h(x;\delta) \equiv \frac{x\xi(x;\delta)}{1-\Xi(x;\delta)}$$

is increasing in x for all $\delta \in [0,1]$ and decreasing in δ for a subset $[0, \widehat{x}(\delta)]$ that may depend on δ.

Assumption (i) is needed to produce the analytical results below. Assumptions (ii) and (iii) impose that the distribution of solvent entrepreneurs faced by smaller banks represent a mean preserving spread with respect that of larger banks. So, apart from indexing the size of the bank, δ is also a measure of diversification.[7] Assumption (iv) is needed for banks' default risk to be reduced with size. This condition is shared by a large family of distributions including their truncations.[8]

In this economy, banks are a payment technology. Take, again, a general bank of size δ. By the time this bank starts functioning at $t = 0$, it has collected, *per location* it serves, $Q(\delta)+D(\delta)$ of assets from investors from which $D(\delta)$ are in the form of deposits and $Q(\delta)$ are in the form of equity. Banks take these funding choices as given. At this point, the bank makes loans, $L(\delta)$, to all the managers on each of the locations that bank serves. These loans are then transferred to workers as wage payments. Workers maintain the funds as deposits until $t = 1$. When providing loans, banks need to satisfy the capital requirement

$$Q(\delta) \geq \kappa L(\delta),$$

where $0 < \kappa < 1$ is a parameter determined by regulation.

With this environment, conditional on the realization of $x(\delta)$, the expost net worth, per location, of a solvent bank, would be

$$x(\delta)\left[1+i^l(\delta)\right]L(\delta) + Q(\delta) - \left[1+i^d(\delta)\right]\left[D(\delta)+L(\delta)\right] \geq 0$$

where $i^l(\delta)$ is the interest rate on loans and $i^d(\delta)$ is the interest rate on deposits. Thus, the bank will be solvent as long as the loan recovery rate satisfies

$$x(\delta) \geq \frac{\left[1+i^d(\delta)\right]\left[D(\delta)+L(\delta)\right] - Q(\delta)}{\left[1+i^d(\delta)\right]L(\delta)} \equiv \underline{x}(\delta). \tag{1}$$

[7] The Appendix shows how this can be achieved with a continuum of locations. The basic idea is to impose that the fraction of solvent entrepreneurs on each location is correlated with those of neighbor points with this correlation decreasing with distance. As banks serve contiguous locations, the average solvency rate among them does not necessarily collapse to the mean μ.

[8] See Bagnoli and Bergstrom [5] as well as Banciu and Mirchandani [6].

Expression (1) defines the minimum loan recovery rate, $\underline{x}(\delta)$, to ensure solvency of the bank. As all banks with the same size start identical and face the same prices, the mass of insolvent banks among those with size δ will be $\Xi\left[\underline{x}(\delta);\delta\right]$.

Then, the expected net worth, per location, of a solvent bank will be[9]

$$\int_{\underline{x}}^{1}\left[x(1+i^l)L+Q-(1+i^d)(D+L)\right]d\Xi(x;\delta).$$

Using the definition of $\underline{x}$ in (1), the expected net worth becomes

$$(1+i^l)L\left\{\int_{\underline{x}}^{1}xd\Xi(x;\delta)-\left[1-\Xi(\underline{x};\delta)\right]\underline{x}\right\}\equiv(1+i^l)L\left[\mu-F(\underline{x};\delta)\right],$$

where μ is the mean of the distribution $\Xi(x;\delta)$, which, under Assumption 1(ii), is independent on δ,

$$F(\underline{x};\delta)=\underline{x}[1-\Xi(\underline{x};\delta)]+G(\underline{x};\delta),$$

and

$$G(\underline{x};\delta)=\int_{0}^{\underline{x}}xd\Xi(x;\delta).$$

As in Bernanke et al. [7], $0\leq F(\underline{x};\delta)\leq 1$. Furthermore,

$$F'(\underline{x};\delta)=1-\Xi(\underline{x};\delta)>0;\ F''(\underline{x};\delta)=-\xi(\underline{x};\delta)<0$$

$$\lim_{\underline{x}\to 0}F(\underline{x};\delta)=0;\ \lim_{\underline{x}\to 1}F(\underline{x};\delta)=\mu$$

$$\lim_{\underline{x}\to 0}G(\underline{x};\delta)=0;\ \lim_{\underline{x}\to 1}G(\underline{x};\delta)=\mu.$$

Depositors will get back their whole deposits only if the bank serving their location is solvent. Thus, they face default risk. Because they are risk averse, a deposit insurance (DI) scheme is introduced in the economy. This DI insures the face value of deposits (not the interest) and is the one negotiating the terms of the deposit contracts with the banks. The details of the relation between the DI and the depositors will be spelled out below. When bargaining the terms of the deposit contract, I assume that the realization of $x(\delta)$ is private information of investors (as equity holders) and banks and not known either by the DI or the depositors (workers). The DI can learn the true value of $x(\delta)$ by hiring auditors. The nominal cost of the audit is born by the depositors and is equal to a proportion $0<\phi<1$ of the amount to be audited, that is, $\phi[1+i^l(\delta)]L(\delta)$. This cost is paid for with deposits. As in Carlstrom and Fuerst [8] or Bernanke et al. [7], this information asymmetry creates a moral hazard problem as banks have incentives to missreport the fraction of loans defaulting. The optimal financial contract should be structured so as to induce the bankers to truthfully

[9] From now on, to simplify notation, I omit reference on variables to the size of the bank, δ, and just maintain it on distributions. Keep in mind, however, that all endogenous variables at the bank level will depend on its size δ.

report the realization of $x(\delta)$. Carlstrom and Fuerst [8] and Bernanke et al. [7] use the findings of Gale and Hellwig [19], Townsend [31] and Williamson [34] to show that in this setup the optimal contract has the form of risky debt. Notice the optimal contract is risky debt in the relation between banks and the DI but has the form of insured deposits from the point of view of depositors. Then, the expected revenues for the DI, per location, will be

$$\int_0^{\underline{x}} \left[x(1+i^l)L+Q\right] d\Xi(x;\delta) - \Xi(\underline{x};\delta)\phi(1+i^l)L + [1-\Xi(\underline{x};\delta)](1+i^d)(D+L).$$

Again, using the definition of $\underline{x}$, these revenues become

$$Q+(1+i^l)L\left[F(\underline{x};\delta)-\phi G(\underline{x};\delta)\right].$$

In this case, notice that

$$\lim_{\underline{x}\to 0}\left[F(\underline{x};\delta)-\phi G(\underline{x};\delta)\right]=0;\ \lim_{\underline{x}\to 1}\left[F(\underline{x};\delta)-\phi G(\underline{x};\delta)\right]=\mu(1-\phi)$$

and

$$F'(\underline{x};\delta)-\phi G'(\underline{x};\delta)=1-\Xi(\underline{x};\delta)-\phi\underline{x}\xi(\underline{x};\delta)=\left[1-\Xi(\underline{x};\delta)\right]\left[1-\phi h(\underline{x};\delta)\right]$$

with $h(\underline{x};\delta)$ defined as

$$h(\underline{x};\delta)=\frac{\underline{x}\xi(\underline{x};\delta)}{1-\Xi(\underline{x};\delta)}.$$

Thus, $\mu - F(\underline{x};\delta)$ represents the fraction of expected bank revenues going to equity holders, $F(\underline{x};\delta)-\phi G(\underline{x};\delta)$ represents the fraction of expected bank revenues going to depositors and $\phi G(\underline{x};\delta)$ represent the fraction of expected bank revenues remunerating auditors. Notice that

$$[\mu-F(\underline{x};\delta)]+[F(\underline{x};\delta)-\phi G(\underline{x};\delta)]+\phi G(\underline{x};\delta)=\mu,$$

i.e. the expected fraction of loans being repaid.

As locations are exante identical, we can define the optimization problem of each bank for a representative location among the ones it serves. Banks take Q, D, i^o, i^l and i^d as given and choose loan supply, L, and the threshold value, $\underline{x}$, to maximize the return for shareholders (investors) over their opportunity cost

$$\frac{\left(1+i^l\right)L\left[\mu-F(\underline{x};\delta)\right]}{\left(1+i^o\right)Q},$$

subject to the participation constraint for depositors

$$Q+(1+i^l)L\left[F(\underline{x};\delta)-\phi G(\underline{x};\delta)\right]\geq\left(1+i^o\right)(L+D), \tag{2}$$

and the capital requirement

$$Q\geq\kappa L. \tag{3}$$

Notice that I am assuming the DI to have access to the certificates of deposits issued by the central bank. Thus, the official rate i^o is the opportunity cost of deposits. On the other hand, because investors are diversifying their equity holdings across all banks in the economy they will not be facing any risk from this investment. Therefore, as owners of the banks they invest in, they want to maximize the return they obtain in each bank over the other investment possibilities, namely, central bank certificates of deposits and bank deposits, both of which provide a return equal to i^o.

When solving this maximization problem, I will not include the constraint (3). After solving the problem I will check whether this constraint binds or not. Let λ be the leverage of the bank defined as

$$\lambda = \frac{L}{Q}.$$

In equilibrium the return on internal funds (equity) at the bank will weakly exceed the opportunity cost. Then, investors will neither deposit any of their initial funds ($D = 0$) nor buy central bank certificates of deposits ($O = 0$) and will provide only equity to the bank. With this in mind, the problem of the bank can be rewritten as choosing leverage λ and the threshold $\underline{x}$ to maximize the equity premium

$$\left(\frac{1+i^l}{1+i^o}\right)[\mu - F(\underline{x};\delta)]\,\lambda, \tag{4}$$

subject to the participation constraint for the DI

$$\left[(1+i^o) - (1+i^l)\,(F(\underline{x};\delta) - \phi G(\underline{x};\delta))\right]\lambda = 1. \tag{5}$$

Let η be the Lagrange multiplier associated with (5). The first order conditions determining bank leverage, λ, the bank solvency threshold level, $\underline{x}$, and the Lagrange multiplier, η, are, respectively,

$$\left(\frac{1+i^l}{1+i^o}\right)[\mu - F(\underline{x};\delta)] + \eta\left[(1+i^l)\,(F(\underline{x};\delta) - \phi G(\underline{x};\delta)) - 1 - i^o\right] = 0 \tag{6}$$

$$F'(\underline{x}) = \eta(1+i^o)\left[F'(\underline{x};\delta) - \phi G'(\underline{x};\delta)\right], \tag{7}$$

and

$$(1+i^o)\lambda = \frac{1}{1 - \left(\frac{1+i^l}{1+i^o}\right)[F(\underline{x};\delta) - \phi G(\underline{x};\delta)]}. \tag{8}$$

As shown in the Appendix, this system defines a positive relationship between the cutoff value, $\underline{x}$, bank leverage, λ, the lagrange multiplier, η, and the loan rate, i^l, for a given level of the official rate $(1+i^o)$. Furthermore, this relationship determines, for a given level of initial equity, Q, a positively sloped loan supply function. Larger loan rates i^l induce banks to increase leverage through providing more loans to managers. The Appendix also shows that the loan supply function is perfectly elastic at rate

$$1 + i^l_{\min} = \frac{1 + i^o}{\mu},$$

while it becomes unbounded as $1 + i^l \to (1 + i^o)/[\mu(1 - \phi)]$. Finally, it can be shown (see the Appendix) that the expected bank equity premium equals

$$\frac{R^e}{1 + i^o} = \left(\frac{1 + i^l}{1 + i^o}\right) [\mu - F(\underline{x}; \delta)] \, \lambda = \frac{1}{(1 + i^o) \, [1 - \phi h(\underline{x}; \delta)]} = \eta \tag{9}$$

while the deposit rate is

$$i^d = i^o + (1 + i^l) \int_{x_{\min}}^{\underline{x}} [\underline{x} - (1 - \phi)x] \, d\Xi(x; \delta). \tag{10}$$

The capital constraint (3) imposes an upper bound on leverage

$$\lambda \leq \frac{1}{\kappa},$$

and, therefore, for a given level of equity Q, there is a limit on the amount of loans banks can provide, L. Notice this leverage bound is common to all banks. To incorporate this constraint use (5) so that it must be the case that

$$\kappa \leq (1 + i^o) - (1 + i^l) \left(F(\underline{x}; \delta) - \phi G(\underline{x}; \delta)\right). \tag{11}$$

In any case, because δ affects the functions $F(\underline{x}; \delta)$ and $G(\underline{x}; \delta)$, banks with different sizes will produce different mappings between $\underline{x}$, i^l, λ and η. This means the constraint (3), or its equivalent (11), may not be binding for some banks.

2.3.3 The deposit insurance scheme

The DI stands between banks and households, negotiating the terms of the deposit contract with banks and providing insurance to depositors. The DI scheme is as follows. The DI insures the face value of deposits (equal to the nominal wage rate, W, per depositor). It collects all deposit revenues from all banks, independently of size, and returns W to each depositor. Any excess of the deposit revenues from banks with respect to the insurance claims from depositors is transferred to *all* depositors in the economy in a lump sum fashion. Let this nominal lump sum transfer be T.

On the one hand, notice that at $t = 0$ agents know how many workers will be hired as auditors and, therefore, will not be producing at $t = 1$. As mentioned above, the total bill to pay auditors is

$$\phi(1 + i^l(\delta))LG(\underline{x}; \delta).$$

With the nominal wage rate being W, the number of auditors to be hired to monitor a bank of size δ in case it goes bankrupt should therefore be

$$N^a(\delta) = \frac{\phi(1+i^l(\delta))L(\delta)G(\underline{x};\delta)}{W}.$$

On the other hand, at $t = 1$ workers receive a total revenue, per location, of

$$Q(\delta) + (1+i^l(\delta))L(\delta)\left[F(\underline{x};\delta) - \phi G(\underline{x};\delta)\right]$$

of which $L(\delta)$ is the wage bill and

$$T(\delta) = Q(\delta) + (1+i^l(\delta))L(\delta)\left[F(\underline{x};\delta) - \phi G(\underline{x};\delta)\right] - L(\delta)$$

is the lump sum interest transfer.

2.4 The nominal side (ii): liquidity risk

As developed in the previous section, a bank of size δ will produce, per location, loans in the amount $L(\delta)$. These loans are initially held by entrepreneurs as deposits to be transferred to the workers they hire. If all workers lived in the very same section of the circle the bank serves (which is where the entrepreneurs are located), the transfer of deposits between entrepreneurs and workers will just be a renaming of these bank's liabilities. However, with an islandwide labor market, entrepreneurs can hire workers from any part of the circle. Because there is a continuum of both workers and entrepreneurs, under the Law of Large Numbers, on average, a fraction δ of those workers will live in the section of the circle served by the bank and the other fraction $(1-\delta)$ will live on outside locations served by the other bank. This means, when transferring deposits to pay for wages, funds equal to $(1-\delta)\delta L(\delta)$ will leave the bank. This is, a fraction $(1-\delta)$ of the total wage payments of the entrepreneurs, $\delta L(\delta)$.

On the other hand, entrepreneurs in the locations not served by the bank will probably hire workers whose accounts are in that bank. Entrepreneurs in those outside locations will go to the other bank and ask for a loan of size $L(1-\delta)$ to pay for wages (so their total wage bill is $(1-\delta)L(1-\delta)$). So, a fraction δ of those transfers will end up in our segment of the circle. This means an inflow of funds to the bank equal to $\delta(1-\delta)L(1-\delta)$. Notice $L(\delta)$ and $L(1-\delta)$ may be different. In the previous section we have seen these decisions depend on the relative sizes of each bank.

Thus, when considering its liquidity needs, the bank foresees it will have a total outflow of funds equal to $(1-\delta)\delta L(\delta)$ and a total inflow of funds equal to $(1-\delta)\delta L(1-\delta)$. The problem is that the *timing* of these flows is uncertain. To model this uncertainty, assume that payments to workers are done continuously throughout the period in a sequential fashion and in a counterclockwise manner starting from an initial point $h \in [0,1]$ within the circle (see, again, Figure 1). This starting point h, though, is random and distributed uniformly on the interval $[0,1]$. Its realization is idiosyncratic to each island. Clearly, if h falls within the locations served by, say, the small bank, as on Figure 1, that bank will start with an inflow of funds as the large bank starts transferring wages to clients of the small bank. On the contrary, if h falls outside the locations served by the small bank, that bank will start with an outflow of funds as it starts transferring wages to clients of the large bank. The reason of modeling the continuous of islands aims at having different banks with different sizes facing different liquidity needs once all idiosyncratic liquidity risks are realized.

All payments among banks are done in reserves. These reserves are provided by the central bank through an open market operation (OMO) performed at the very beginning of period $t = 0$, before uncertainty about the initial location receiving funds, h, is resolved. The OMO is done through a fixed rate, full allotment, tender. The interest rate of the operation is the official rate i^o. Let $M(\delta)$ denote the allotment to bank δ. Once h is resolved, given this allotment to bank $\delta = \{\delta_S, \delta_L\}$, $M(\delta)$, and its planned loan provision, $L(\delta)$, each bank on each island can compute the evolution of its liquidity throughout period $t = 0$. Without recourse to the interbank market, the bank starts period $t = 0$ with $M(\delta)$ reserves. Now, two things may happen, namely, that the realization of h falls within the locations served by bank δ or within the locations served by the other bank. To fix ideas, I will work out the problem faced by the small bank ($\delta = \delta_S$) in Figure 1. I will then provide equivalent expressions for the large bank of this island which can be worked out similarly.

Start by assuming that the initial location receiving payments is not served by the small bank, i.e. $h \in (0, 1-\delta]$. Because payments are done continuously in a counterclockwise manner, funds will leave the bank at the rate $\delta L(\delta)$ for the first fraction h of the period. Thus, the average reserve holdings for this section of the period will be

$$\int_0^h [M(\delta) - \delta L(\delta) t]\, dt. \tag{12}$$

Once these payments are done, it will be the turn of the locations served by the small bank which means this bank will experience an inflow of funds at the rate $(1-\delta)L(1-\delta)$. Notice at the time the bank starts receiving funds, its liquidity position has reached a level of $M(\delta) - \delta L(\delta)h$. Thus, the average reserve holdings for this section of the period will be

$$\int_0^\delta [M(\delta) - \delta L(\delta)h + (1-\delta)L(1-\delta)t]\, dt. \tag{13}$$

Once all locations of the bank have received payment, it will be the turn of the reminder outside locations. They represent a mass $1-\delta-h$ of the circle and, again, funds will leave the bank at the rate $\delta L(\delta)$. Notice that the level of funds at the bank starts at $M(\delta) - \delta L(\delta)h + (1-\delta)L(1-\delta)\delta$ so that the average reserve holdings will be

$$\int_0^{1-\delta-h} [M(\delta) - \delta L(\delta)h + (1-\delta)L(1-\delta)\delta - \delta L(\delta)t]\, dt. \tag{14}$$

Adding terms (12) through (14), the average reserve position of the bank throughout period $t = 0$ would be

$$M(\delta) - h\delta\mathbb{L} + \frac{\delta(1-\delta)}{2}\mathbb{L} + \frac{\delta(1-\delta)}{2}\left[L(1-\delta) - L(\delta)\right] - R(\delta) \tag{15}$$

where

$$\mathbb{L} = (1-\delta)L(1-\delta) + \delta L(\delta)$$

is total payments done at the island and $R(\delta)$ is the reserve requirement that needs to be satisfied by the bank (which are predetermined and may equal 0).

On the other hand, if the initial location receiving payments is served by the small bank, i.e. $h \in (1-\delta, 1]$, then funds will initially enter the bank at the rate $(1-\delta)L(1-\delta)$ for the first fraction $h - 1 + \delta$ of the period. Performing similar computations as above produces an average reserve position of the bank equal to

$$M(\delta) - (1-h)(1-\delta)\mathbb{L} + \frac{\delta(1-\delta)}{2}\mathbb{L} + \frac{\delta(1-\delta)}{2}\left[L(1-\delta) - L(\delta)\right] - R(\delta). \tag{16}$$

By inspecting (15) one can see that, given the OMO allotment, $M(\delta)$, and the planned loan and deposit productions, $L(\delta)$ and $L(1-\delta)$, whenever h falls on outside locations, so that $h \in (0, 1-\delta]$, the larger the realization of h the more likely is the average reserve position to be negative. In particular, for

$$0 \leq h \leq h_A(\delta) \equiv \frac{M(\delta) - R(\delta)}{\delta\mathbb{L}} + \frac{(1-\delta)}{2} + \left(\frac{1-\delta}{2}\right)\left(\frac{L(1-\delta) - L(\delta)}{\mathbb{L}}\right) \tag{17}$$

the average reserve position will be positive while it will be negative for

$$h_A(\delta) < h \leq 1 - \delta. \tag{18}$$

Similarly, by inspecting (16) one can see that, given the OMO allotment, $M(\delta)$, and the planned loan and deposit productions, $L(\delta)$ and $L(1-\delta)$, whenever h falls on inside locations, so that $h \in (1-\delta, 1]$, the smaller the realization of h the more likely is the average reserve position to be negative. In particular, for

$$1 - \delta \leq h \leq h_B(\delta) \equiv 1 - \frac{M(\delta) - R(\delta)}{(1-\delta)\mathbb{L}} - \frac{\delta}{2} - \frac{\delta}{2}\left(\frac{L(1-\delta) - L(\delta)}{\mathbb{L}}\right) \tag{19}$$

the average reserve position will be negative while it will be positive for

$$h_B(\delta) < h \leq 1. \tag{20}$$

Assume the central bank offers two permanent standing facilities. First, there is a marginal lending facility in which commercial banks can obtain as much liquidity as needed at a penalty rate $i^{LF} > i^o$.[10] Second, there is a

[10] Actually, when in place, commercial banks can borrow from marginal lending facilities against the presentation of eligible collateral as they do with the OMO.

deposit facility in which commercial banks can place any excess liquidity remunerated at the rate $i^{DF} < i^{o}$.[11] Now assume an economywide interbank market opens at the beginning of period $t = 0$ right after the uncertainty about h is resolved. In this market, banks with excess reserves can provide liquidity to banks with liquidity deficits. Let $I(\delta, h)$ be the lending of the bank in the interbank market (borrowing if negative) for a particular realization of h. Clearly, under no uncertainty about the liquidity position at $t = 1$, which is the case here once the value of h is known, if the return on lending in the interbank market is larger than the rate of the deposit facility, i^{DF}, and the cost of borrowing in the interbank market is smaller than the rate of the marginal lending facility, i^{LF}, it must be the case that the supply of funds in the interbank market will equal either (15) or (16), that is, if $h \in (0, 1-\delta]$

$$I_A(\delta, h) = M(\delta) - R(\delta) - h\delta\mathbb{L} + \frac{\delta(1-\delta)}{2}\mathbb{L} + \frac{\delta(1-\delta)}{2}\left[L(1-\delta) - L(\delta)\right] \quad (21)$$

while if $h \in (1-\delta, 1]$

$$\begin{aligned} I_B(\delta, h) &= M(\delta) - R(\delta) - (1-h)(1-\delta)\mathbb{L} + \frac{\delta(1-\delta)}{2}\mathbb{L} \\ &\quad + \frac{\delta(1-\delta)}{2}\left[L(1-\delta) - L(\delta)\right]. \end{aligned} \quad (22)$$

Notice a particular small bank in a particular island will be a lender or a borrower in the interbank market depending on its particular realization of h as compared with its threshold (17) and (19).

Because there are solvency risks in this economy and because loans in the interbank market are unsecured, a bank demanding reserves will have to pay a spread or risk premium, call it $s(\delta)$, to lenders. With this premium, lenders are sure to obtain on average a rate i on that loan. Since the default probability of a bank of size δ is $\Xi\left[\underline{x}(\delta); \delta\right]$, the spread paid by these banks in case they borrow in the interbank market should satisfy

$$1 + i = \left[1 - \Xi\left(\underline{x}(\delta); \delta\right)\right]\left[1 + i + s(\delta)\right]$$

or

$$s(\delta) = \frac{\Xi\left[\underline{x}(\delta); \delta\right]}{1 - \Xi\left[\underline{x}(\delta); \delta\right]}(1 + i). \quad (23)$$

Thus, the condition for trade in the interbank market to happen is $i^{DF} < i < i + s(\delta) < i^{LF}$.

With all this information, we are now ready to solve for the bank's bid at the OMO. Remember by the time this OMO takes place banks are uncertain about

[11] Nowdays many central banks offer explicitly such facilities with the aim of setting a corridor for overnight rates. Such corridors do also implicitly exist whenever these facilities are not provided. In such cases, the rate of the discount window (plus a valuation of any nonpecuniary costs associated with its use) can provide an upper bound (corresponding to the marginal lending rate) while zero provides the lower bound (corresponding to the deposit rate).

the timing of payments, namely, the realization of h. Thus, the bank maximizes the expected profits from liquidity management. The costs of obtaining liquidity are represented by the official rate $1 + i^o$. The net revenues will depend on the position of the bank in the interbank market which itself will depend upon the realization oh h. For $h \in (0, h_A]$ and $h \in (h_B, 1]$ the bank will have excess reserves that can be loaned out in the interbank market obtaining a rate $1 + i$ while for $h \in (h_A, 1 - \delta]$ or $h \in (1 - \delta, h_B]$ the bank will have a reserve deficit that will be borrowed from the interbank market at the rate $1 + i + s(\delta)$. Then expected profits from liquidity management equal

$$\begin{aligned} \Pi\left[M(\delta)\right] &= (1+i)\int_0^{h_A(\delta)} I_A(\delta,h)dh + (1+i+s(\delta))\int_{h_A(\delta)}^{1-\delta} I_A(\delta,h)dh \\ &\quad +(1+i+s(\delta))\int_{1-\delta}^{h_B(\delta)} I_B(\delta,h)dh + (1+i)\int_{h_B(\delta)}^{1} I_B(\delta,h)dh \\ &\quad -(1+i^o)M(\delta) \end{aligned} \tag{24}$$

where $I_A(\delta, h)$ and $I_B(\delta, h)$ are, respectively, displayed in (21) and (22).

Taking first order conditions with respect to $M(\delta)$ in (24) produces the following demand for reserves at the OMO

$$\begin{aligned} M(\delta) &= R(\delta) + \frac{\delta(1-\delta)}{2}\mathbb{L} - \frac{\delta(1-\delta)}{2}\left[L(1-\delta) - L(\delta)\right] \\ &\quad -\delta(1-\delta)\mathbb{L}\left(\frac{i^o - i}{s(\delta)}\right). \end{aligned} \tag{25}$$

Substituting this demand in either (21) or (22) yields the supply of funds in the interbank market: if $h \in (0, 1 - \delta]$

$$I_A(\delta,h) = \delta(1-\delta)\mathbb{L}\left(1 - \frac{i^o - i}{s(\delta)}\right) - h\delta\mathbb{L} \tag{26}$$

while if $h \in (1 - \delta, 1]$

$$I_B(\delta,h) = \delta(1-\delta)\mathbb{L}\left(1 - \frac{i^o - i}{s(\delta)}\right) - (1-h)(1-\delta)\mathbb{L}. \tag{27}$$

Several comments regarding (25), (26) and (27) are in order here. First, notice the role of reserve requirements. They are just incorporated deterministically in the reserve demand at the OMO and do not affect the supply of funds in the interbank market. These requirements could well be set at 0 without no implications on the equilibrium at the interbank market. Second, because each island contains a small bank and because the realization of h is idiosyncratic to each island, there will be one small bank on each value of h and the total supply of funds in the interbank market by small banks, $I(\delta)$, will be

$$\begin{aligned} I(\delta) &= \int_0^{1-\delta} I_A(\delta,h)dh + \int_{1-\delta}^{1} I_B(\delta,h)dh \\ &= \delta(1-\delta)\left(\frac{1}{2} - \frac{i^o - i}{s(\delta)}\right)\mathbb{L}. \end{aligned} \tag{28}$$

Thus, the only individual variable entering the supply of interbank funds by small banks is their borrowing spread. Again, all deterministic reserve needs derived from different loan behavior or different sizes are anticipated and taken care of with the bidding behavior at the OMO.

From the discussion above, we can see that although the derivations have been done for small banks, nothing prevents the same computations to be done for the large bank. The only difference is that if $h \in (0, 1-\delta]$ the bank will start with an inflow of funds while if $h \in (1-\delta, 1]$ the bank starts with an outflow of funds. So, repeating the derivations above for the large bank, the supply of funds in the interbank market: if $h \in (0, 1-\delta]$ would be[12]

$$I_A(1-\delta, h) = h\delta\mathbb{L} - \delta(1-\delta)\mathbb{L}\frac{i^o - i}{s(1-\delta)}.$$

The thresholds for this magnitude to be positive or negative is now

$$\begin{aligned} h_A(1-\delta) \quad &\equiv \quad \frac{(1-\delta)}{2} + \left(\frac{1-\delta}{2}\right)\left(\frac{L(1-\delta) - L(\delta)}{\mathbb{L}}\right) \\ &\quad - \frac{M(1-\delta) - R(1-\delta)}{(1-\delta)\mathbb{L}} \end{aligned} \tag{29}$$

so that the reserve position will be negative if

$$0 < h \leq h_A(1-\delta)$$

and positive if

$$h_A(1-\delta) < h \leq 1-\delta.$$

On the other hand, if $h \in (1-\delta, 1]$, the supply of funds is

$$I_B(1-\delta, h) = (1-h)(1-\delta)\mathbb{L} - \delta(1-\delta)\mathbb{L}\frac{i^o - i}{s(1-\delta)}.$$

The thresholds for this magnitude to be positive or negative is now

$$h_B(1-\delta) \equiv 1 + \frac{M(1-\delta) - R(1-\delta)}{\delta\mathbb{L}} - \frac{\delta}{2} - \frac{\delta}{2}\left(\frac{L(1-\delta) - L(\delta)}{\mathbb{L}}\right) - \tag{30}$$

so that the reserve position will be positive if

$$1-\delta < h \leq h_B(1-\delta)$$

and negative if

$$h_B(1-\delta) < h \leq 1.$$

Similarly, the demand for funds at the OMO for large banks will be,

$$\begin{aligned} M(1-\delta) \quad &= \quad R(1-\delta) + \frac{\delta(1-\delta)}{2}\mathbb{L} + \frac{\delta(1-\delta)}{2}\left[L(1-\delta) - L(\delta)\right] \\ &\quad -\delta(1-\delta)\mathbb{L}\left(\frac{i^o - i}{s(1-\delta)}\right). \end{aligned} \tag{31}$$

[12] Notice in this case the large bank will start with an inflow of funds.

Again, as there is also a continuum of islands with one bank each and the realization of h is idiosyncratic to each island, the total supply of funds in the interbank market by large banks, $I(1-\delta)$, will be

$$\begin{aligned} I(1-\delta) &= \int_0^{1-\delta} I_A(1-\delta,h)dh + \int_{1-\delta}^{1} I_B(1-\delta,h)dh \\ &= \delta(1-\delta)\left(\frac{1}{2}-\frac{i^o-i}{s(1-\delta)}\right)\mathbb{L}. \end{aligned} \tag{32}$$

Looking at (28) and (32) one can see that the net position in the interbank market of large and small banks will depend on the spread between the official and the interbank rates, $i^o - i$, together with the borrowing premium of each class of banks, $s(\delta)$ and $s(1-\delta)$. Below I will solve for the interbank rate and evaluate the net position of large and small banks in the interbank market.

There is an additional point to be made here. As the reader may have noticed, liquidity costs are not included in the solvency problem above. One may think that the revenues/costs of maintaining a liquidity position should be taken into account when determining the amount of loans to be made by the bank. This is not done for two reasons. First, the liquidity and solvency problems are separated for tractability. The solvency problem becomes complicated if it includes also the liquidity problem. This is because the liquidity costs, affecting solvency risk, depend on the default probability through the interbank borrowing premium. Second, actual depository institutions take liquidity management as a residual of other operations of the bank. This means they conduct operations on other parts of the balance sheet and then find reserves to back up those positions. In this sense, here I assume these liquidity costs of interbank trade associated with payments by banks are relatively small and therefore negligible when considering solvency risks.

Finally, the reader may have also noticed that on $t = 1$, as workers pay entrepreneurs for the purchase of consumption goods, new transfers of reserves between banks will take place depending on the consumption patterns of these workers. On this regard, I concentrate only on the net positions of banks in the interbank at $t = 0$ and ignore flows on $t = 1$ by assuming that the interbank market is closed on this later date and reserve compensation is done through the standing facilities of the central bank. This is done for two reasons. First, I would like to keep the analysis as simple as possible without the need to explicitly modelling the consumption patterns of workers through different locations or the resolution of bank's defaults. Second, in a fully dynamic general equilibrium model, there will be a continuation value of keeping reserves at $t = 1$ if the time horizon continues after that date. As I have not modelled this continuation value, I prefer to ignore reserve demand in the interbank market on that day and concentrate on net positions at $t = 0$.

2.5 The real side: workers and entrepreneurs

In each location there is a continuum of measure one of workers and a measure one of entrepreneurs. As these locations are formally identical, I will describe the problem in one of them. Workers supply inelastically one unit of labor to the market. Independently of working as auditors or for an entrepreneur, they all receive the certain equilibrium nominal wage W plus the DI transfer T so that their utility will be

$$u\left(\frac{W+T}{P}\right) = u(w+t) \tag{33}$$

where P is the nominal price level, w is the real wage and t is the real value of the DI transfer. In this expression u is a strictly increasing, strictly concave function satisfying the Inada conditions.

Entrepreneurs in any location j demand labor. If an entrepreneur e in location j hires n_{ej} workers, his output will be

$$q_{ej} z (n_{ej})^{\alpha},$$

with $0 < \alpha < 1$. The parameter z indexes the level of technology and is equal for all entrepreneurs in all locations on any island. Furthermore, q_{ej} is an indicator function taking value 1 or 0. Treating all entrepreneurs symmetrically, from the point of view of a single entrepreneur in a particular location j at $t = 0$, he will be able to produce ($q_{ej} = 1$) with probability x_j and will not be able to produce ($q_{ej} = 0$) with probability $1 - x_j$. Therefore, with probability x_j the entrepreneur will have access at $t = 1$ to the real value of the profits derived from selling his production while with probability $1 - x_j$ the entrepreneur will not be able to produce anything. From the point of view of the location, x_j is also the random fraction of entrepreneurs in location j who will be able to produce at $t = 1$. Notice this probability is itself a random variable with marginal distribution function $\Xi(x)$.[13]

Because managers enter the period with no resources, and because wages need to be paid for in advance, managers have to ask for a loan to the bank in their location. Because x_j and q_{ej} are realized after the labor decision is taken and because of limited liability, all entrepreneurs in location j choose the same amount of labor, $n_{ej} = n_j$, to maximize expected real profits

$$E(\pi_j) = \mu \left[z n_j^{\alpha} - (1 + i_j^l) w n_j \right] \tag{34}$$

where i_j^l is the nominal interest rate on loans in the location the entrepreneur lives which equals either $i^l(\delta_S)$ or $i^l(\delta_L)$ on which bank serves that location. The first order condition with respect to labor input for the manager is

$$\alpha z n_j^{\alpha-1} = (1 + i_j^l) w \tag{35}$$

[13]The Appendix shows how to connect the distribution $\Xi(x)$ at the location level with the distribution $\Xi(x, \delta)$ at the bank level.

so that demand for labor is

$$n(i_j^l, w) = \left[\frac{\alpha z}{(1 + i_j^l)w} \right]^{1/(1-\alpha)}, \tag{36}$$

and expected real profits are

$$E(\pi_j) = \mu \left[(1-\alpha) \left(\frac{z\alpha^\alpha}{(1 + i_j^l)^\alpha w^\alpha} \right)^{1/(1-\alpha)} \right] > 0.$$

Nominal loan demand of these entrepreneurs in that location will therefore be

$$L_j^d = Wn(i_j^l, w) = W \left[\frac{\alpha z}{(1 + i_j^l)w} \right]^{1/(1-\alpha)} = P \left[\frac{\alpha z}{(1 + i_j^l)w^\alpha} \right]^{1/(1-\alpha)}. \tag{37}$$

3 General equilibrium

3.1 Characterization

As mentioned above, there are two types of banks in the economy with sizes $0 < \delta_S < 0.5 < \delta_L < 1$. For simplicity, let $\delta s = \delta$ and $\delta_L = 1 - \delta$. To characterize the equilibrium of this economy, start with investors. These agents split their nominal assets, A, between equity in all banks in the circle, that is,

$$\delta Q(\delta) + (1-\delta)Q(1-\delta) = A.$$

This is repeated for all islands in the economy as they all start identical. For all banks to be financed through equity, and therefore function, it must be the case that investors should receive the same expected return on all types of equity. Furthermore, competition in the market for bank equity will drive down the return on equity to its opportunity cost, i^o. Using (9) this means that

$$\eta(\delta) = \eta(1-\delta) = \eta = 1.$$

This condition determines, for each bank, an equilibrium solvency threshold level $\underline{x}(\delta_b)$ satisfying

$$1 - \phi h\left[\underline{x}(\delta_b); \delta_b\right] = \frac{1}{1 + i^o}, \tag{38}$$

where δ_b is either δ or $1 - \delta$. With these threshold levels, FOCs (6) and (8) provide with the equilibrium leverage $\lambda(\delta_b)$ and loan rate $i^l(\delta_b)$ for each bank

$$\begin{aligned} 1 + i^o \quad &= \quad \left[\frac{1 + i^l(\delta_b)}{1 + i^o} \right] \left[\mu - F\left[\underline{x}(\delta_b); \delta_b\right]\right] \\ &\quad + (1 + i^l(\delta_b)) \left(F\left[\underline{x}(\delta_b); \delta_b\right] - \phi G\left[\underline{x}(\delta_b); \delta_b\right] \right) \end{aligned} \tag{39}$$

$$\lambda(\delta_b) = \min\left\{\frac{1}{1+i^o-(1+i^l(\delta_b))\left[F\left[\underline{x}(\delta_b);\delta_b\right]-\phi G\left[\underline{x}(\delta_b);\delta_b\right]\right]}, \frac{1}{\kappa}\right\}. \quad (40)$$

Leverage $\lambda(\delta_b)$ together with bank equity $Q(\delta_b)$ generate the supply of loans to be produced in each of the locations served by each bank

$$L(\delta_b) = \lambda(\delta_b)Q(\delta_b).$$

On the other hand, given the real wage rate, w, and the loan rates $i^l(\delta_b)$, expression (36) determines the demands for labor by entrepreneurs

$$n\left[i^l(\delta_b), w\right] = \left[\frac{\alpha z}{(1+i^l(\delta_b))w}\right]^{1/(1-\alpha)}.$$

Since labor demand is the same for all locations served by each bank, equilibrium in the labor market implies that

$$\begin{aligned} 1-N^a &= \delta n\left[i^l(\delta), w\right] + (1-\delta)n\left[i^l(1-\delta), w\right] \\ &= \delta\left[\frac{\alpha z}{(1+i^l(\delta))w}\right]^{1/(1-\alpha)} + (1-\delta)\left[\frac{\alpha z}{(1+i^l(1-\delta))w}\right]^{1/(1-\alpha)} \end{aligned} \quad (41)$$

where the left hand side is the aggregate labor supply for production purposes which equals, $1-N^a$, that is, total labor minus total auditing services. The total measure of auditors for the whole economy will be

$$\begin{aligned} N^a &= N^a(\delta)\Xi\left[\underline{x}(\delta);\delta\right] + N^a(1-\delta)\Xi\left[\underline{x}(1-\delta);1-\delta\right] \\ &= \frac{\mu(1+i^l(\delta))L(\delta)G\left[\underline{x}(\delta);\delta\right]}{W}\Xi\left[\underline{x}(\delta);\delta\right] \\ &\quad + \frac{\mu(1+i^l(1-\delta))L(1-\delta)G\left[\underline{x}(1-\delta);1-\delta\right]}{W}\Xi\left[\underline{x}(1-\delta);1-\delta\right]. \end{aligned} \quad (42)$$

Additionally, given the price level, P, the real wage rate, w, and the loan rates $i^l(\delta_b)$, expression (37) determines nominal loan demands by entrepreneurs

$$L^d\left[i^l(\delta_b), w, P\right] = Wn\left[i^l(\delta_b), w\right] = P\left[\frac{\alpha z}{(1+i^l(\delta_b))w^\alpha}\right]^{1/(1-\alpha)}.$$

Because loan demand is the same for all locations served by each bank, equilibrium in the loan market implies that

$$L(\delta) = L^d\left[i^l(\delta), w\right] = P\left[\frac{\alpha z}{(1+i^l(\delta))w^\alpha}\right]^{1/(1-\alpha)}$$

and

$$L(1-\delta) = L^d\left[i^l(1-\delta), w\right] = P\left[\frac{\alpha z}{(1+i^l(1-\delta))w^\alpha}\right]^{1/(1-\alpha)}.$$

These expressions together with (41) and (42) determine the price level P together with the nominal (or real) wage rate W (or w). Notice loan supply must equal loan demand on each location, i.e.

$$L^d\left[i^l(\delta_b), w\right] = L(\delta_b) = \lambda(\delta_b)Q(\delta_b).$$

Thus, given the chosen leverages, $\lambda(\delta_b)$, bank capital must be allocated across banks for this condition to be met for all of them. Additionally, deposit rates $i^d(\delta_b)$ are computed from

$$i^d(\delta_b) = i^o + (1 + i^l(\delta_b)) \int_{x_{\min}}^{\underline{x}(\delta_b)} \left[\underline{x}(\delta_b) - (1-\phi)x\right] d\Xi(x;\delta).$$

Because there is a measure 1 of both large and small banks, from (28) and (32) equilibrium in the money market implies

$$I(\delta) + I(1-\delta) = \delta(1-\delta)\left(\frac{1}{2} - \frac{i^o - i}{s(\delta)}\right)\mathbb{L} + \delta(1-\delta)\left(\frac{1}{2} - \frac{i^o - i}{s(1-\delta)}\right)\mathbb{L} = 0$$

so that the equilibrium interbank rate is

$$i = i^o - \frac{s(\delta)s(1-\delta)}{s(\delta)+s(1-\delta)}. \tag{43}$$

Substituting back this rate in (28) and (32) produces the net position of small and large banks:

$$I(\delta) = \frac{\delta(1-\delta)}{2}\left(\frac{s(\delta)-s(1-\delta)}{s(\delta)+s(1-\delta)}\right)\mathbb{L} \tag{44}$$

and

$$I(1-\delta) = \frac{\delta(1-\delta)}{2}\left(\frac{s(1-\delta)-s(\delta)}{s(\delta)+s(1-\delta)}\right)\mathbb{L}. \tag{45}$$

Notice that although the interbank rate is below the official rate, $i < i^o$, see (43), the rate paid by borrowers which includes the risk premium exceeds the official rate, i.e., for small banks,

$$i + s(\delta) = i^o + \frac{s(\delta)^2}{s(\delta)+s(1-\delta)} > i^o$$

while for large banks

$$i + s(1-\delta) = i^o + \frac{s(1-\delta)^2}{s(\delta)+s(1-\delta)} > i^o.$$

Furthermore, substituting the equilibrium rates in (25) and (31) produces, respectively, the bids of small and large banks:

$$\begin{aligned} M(\delta) &= R(\delta) + \frac{\delta(1-\delta)}{2}\mathbb{L} - \frac{\delta(1-\delta)}{2}\left[L(1-\delta) - L(\delta)\right] \\ &\quad -\delta(1-\delta)\mathbb{L}\frac{s(1-\delta)}{s(\delta)+s(1-\delta)} \end{aligned} \tag{46}$$

and

$$M(1-\delta) = R(1-\delta)+\frac{\delta(1-\delta)}{2}\mathbb{L}+\frac{\delta(1-\delta)}{2}\left[L(1-\delta)-L(\delta)\right]$$
$$-\delta(1-\delta)\mathbb{L}\frac{s(\delta)}{s(\delta)+s(1-\delta)}. \tag{47}$$

Finally, for completeness, notice that substituting equilibrium bids (46) and (47) in the thresholds (17), (19), (29) and (30) yields

$$h_A(\delta) = h_A(1-\delta) = (1-\delta)\frac{s(\delta)}{s(\delta)+s(1-\delta)}$$

while

$$h_B(\delta) = h_B(1-\delta) = 1-\delta\frac{s(\delta)}{s(\delta)+s(1-\delta)}.$$

Notice $0 < h_A < 1-\delta < h_B < 1$ as needed.

3.2 The role of diversification, δ

As mentioned above, the system (38)-(40) determines the cutoff value, $\underline{x}(\delta_b)$, bank leverage, $\lambda(\delta_b)$, and the loan rate, $i^l(\delta_b)$, for small and large banks (so that δ_b could be either δ or $1-\delta$). The question is how these equilibrium values change with the size (diversification) of banks, δ_b. Unfortunately, the answer to this question depends on the particular set of distribution function $\Xi(x;\delta_b)$ and there are no general results. To see this, applying the implicit function theorem to (38), we see that

$$\underline{x}'(\delta_b) = -\frac{\partial h\left(\underline{x};\delta_b\right)}{\partial \underline{x}}\left[\frac{\partial h\left(\underline{x};\delta_b\right)}{\partial \delta_b}\right]^{-1}$$

evaluated at the equilibrium threshold $\underline{x}(\delta_b)$. By Assumption 1(iv) the denominator is positive and the numerator is negative provided $\underline{x}(\delta_b) < \widehat{x}(\delta_b)$. This means that, under Assumption 1, the threshold $\underline{x}(\delta_b)$ is increasing with diversification.

However, another issue is the effect of diversification on the bank default probability, $\Xi\left[\underline{x}(\delta_b);\delta_b\right]$, and leverage $\lambda(\delta_b)$. The total effect on the default probability is

$$\frac{d\Xi\left[\underline{x}(\delta_b);\delta_b\right]}{d\delta_b} = \frac{\partial \Xi\left(\underline{x};\delta_b\right)}{\partial \underline{x}}\underline{x}'(\delta_b)+\frac{\partial \Xi\left(\underline{x};\delta_b\right)}{\partial \delta_b},$$

again evaluated at the equilibrium threshold $\underline{x}(\delta_b)$. The first term of the right hand side is positive while, in general, for low values of $\underline{x}(\delta_b)$, the second term is negative. Thus, the total derivative is ambiguous. This makes the effect on leverage ambiguous too.

Intuition, however, tells us that an increase in risk associated with an decrease in diversification, δ_b, would drive an increase in the default probability and a reduction in leverage.[14] The reason is that keeping the threshold $\underline{x}(\delta_b)$ and leverage, $\lambda(\delta_b)$, constant, an increase in risk will make some banks raise their returns while some other banks will experiment larger losses. Therefore, If small banks had the same threshold, leverage and loan rates than large banks, they will produce larger monitoring costs reducing the return of depositors and their equity holders would have larger returns. To drive the returns of both equity holders and depositors to their opportunity costs, bank leverage have to decrease as well as the threshold. The reduction on the threshold, though, does not totally compensate the increase in default risk so that small banks default at a larger rate than large banks. Then, we should observe

$$\underline{x}(\delta) < \underline{x}(1-\delta);\ \Xi\left[\underline{x}(\delta);\delta\right] > \Xi\left[\underline{x}(1-\delta);1-\delta\right];\ \text{and } \lambda(\delta) < \lambda(1-\delta).$$

If this is the case, then expression (23) tells us that the risk premium small banks have to pay to borrow at the interbank market should be larger than that of large bank, that is

$$s(\delta) > s(1-\delta)$$

so that, from (44) and (45), small banks are net lenders in the money market ($I(\delta) > 0$) while large banks are net net borrowers ($I(1-\delta) < 0$). Furthermore, let $\Theta(\delta_b)$ be the excess reserves above reserve requirements and expected payment needs bank $\delta_b = \{\delta, 1-\delta\}$ bids at the OMO, that is

$$\Theta(\delta_b) = M(\delta_b) - R(\delta_b) - \delta_b(1-\delta_b)\left[L(\delta_b) - L(1-\delta_b)\right].$$

Looking at (46) and (47), these excess reserves are

$$\begin{aligned}\Theta(\delta_b) \quad = \quad & \frac{\delta_b(1-\delta_b)}{2}\mathbb{L} - \frac{\delta_b(1-\delta_b)}{2}\left[L(\delta_b) - L(1-\delta_b)\right] \\ & -\delta_b(1-\delta_b)\mathbb{L}\frac{s(1-\delta_b)}{s(\delta_b)+s(1-\delta_b)}.\end{aligned}$$

It is easy to see that lower payment needs, $L(\delta) < L(1-\delta)$, by small banks together with a larger premium, $s(\delta) > s(1-\delta)$, make small banks demand a larger amount of excess reserves at the OMO as compared with large banks. These excess reserves are then offered at the interbank market.

4 Conclusions

This paper develops a new model of banking and payment system to rationalize the idea that small banks tend to be net sellers of funds in money markets while large banks tend to be net borrowers. When developing the model, care has been taken in reproducing actual institutions present in our monetary systems. In particular, broad money (deposits) is created by commercial banks and used by

[14] See Acosta-Ormaechea and Morozumi [1], Cesa-Bianchi and Fernandez-Corugedo [9], Chugh [12] or Dorofeenko et al. [17], among others, for a similar discussion within the context of an agency cost model between banks and entrepreneurs.

the nonfinancial sector of the economy to finance the real activity. At the same time, narrow money is created by the central bank, it is used by commercial banks to face the net payments derived from the creation of broad money and exchange between these banks through the interbank market. The connection between these two layers of liquidity (narrow and broad) is the endogenous loan and deposit creation by commercial banks.

With this model in hand, the paper then shows that the reason for the money market dichotomy associated with bank size can be linked to the lower diversification and higher solvency risks exhibited by small banks. Because of higher risk, small banks face higher premiums when borrowing unsecured from money markets. Thus, they have incentives to bid an excess of reserves at the central bank open market operation and offer these reserves at the interbank market. Large banks face the opposite problem. There is no need to impose other types of exogenous asymmetries between banks with different sizes with respect to the degree of risk aversion, deposit-taking costs or opacity. Of course these other dimensions could be relevant to explain the observed differences between large and small banks. However, this paper contends that the mechanism associated with these differences should be analyzed in a model as presented here that reproduces actual monetary institutions.

The model presented here produces several testable implications. Apart from the money market dichotomy, small banks in the model maintain lower leverage ratios, are bankrupt at a higher rate, or pay larger borrowing rates in money markets than larger banks. Also, the model is static and there is no aggregate risk. Thus, it is just natural to embed it into a DSGE model to analyze the dynamic effects associated with endogenous leverage at the banking sector. All these questions are left for future research.

References

[1] Acosta-Ormaechea, S., and A. Morozumi (2012): "Idiosyncratic uncertainty, asymmetric information, and private credit" CFCM Working Paper 12/12.

[2] Allen, L., S. Peristiani, and A. Saunders (1989): "Bank size, collateral, and net purchase behavior in the federal funds market: Empirical evidence" *Journal of Business* 62(4), pp. 501-515.

[3] Allen, L., and A. Saunders (1986): "The large-small bank dichotomy in the federal funds market" *Journal of Banking & Finance* 10 (2), pp. 219-230.

[4] Ashcraft, A., J. McAndrews, and D. Skeie (2011): "Precautionary reserves and the interbank market" *Journal of Money, Credit and Banking* 43(7), pp. 311-348.

[5] Bagnoli, M, and T. Bergstrom (2005): "Log-concave probability and its applications", *Economic Theory* 26(2), pp. 445–469.

[6] Banciu, M. and P. Mirchandany (2013): "Technical note - New results concerning probability distributions with increasing generalized failure rates" *Operations Research* 61(4), pp. 925–931.

[7] Bernanke, B., M. Gertler, and S. Gilchrist (1999): "The financial accelerator in a quantitative business cycle framework", in J. B. Taylor and M. Woodford, editors, *Handbook of Macroeconomics*, Amsterdam, 1999. Elsevier Science.

[8] Carlstrom, C., and T. Fuerst (1997): "Agency costs, net worth, and business fluctuations: a computable general equilibrium analysis", *American Economic Review* 87, pp. 893-910.

[9] Cesa-Bianchi, A. and E. Fernandez-Corugedo (2014): "Uncertainty in a model with credit frictions", Bank of England Working Paper No. 496.

[10] Chari, V.V. and C. Phelan (2014): "On the social usefulness of fractional reserve banking" *Journal of Monetary Economics* 65, pp. 1-13.

[11] Chen, A. H., and S. C. Mazumdar (1992): "An instantaneous control model of bank reserves and federal funds management" *Journal of Banking and Finance* 16(6), pp. 1073-1095.

[12] Chugh, S. K. (2016): "Firm risk and leverage-based business cycles" *Review of Economic Dynamics* 20, pp.111–131.

[13] Cocco, J. F., F. J. Gomes, N. C. Martins (2009): "Lending relationships in the interbank market" *Journal of Financial Intermediation* 18, pp. 24–48.

[14] Constancio, V. (2011): "Challenges to monetary policy in 2012", speech at the 26th International Conference on Interest Rates, available at http://www.ecb.europa.eu/press/key/date/2011/html/sp111208.en.html.

[15] Craig, B., and G. von Peter (2010): "Interbank tiering and money center banks" BIS working paper 320.

[16] Disyatat, P. (2011): "The bank lending channel revisited" *Journal of Money, Credit and Banking* 43 (4), pp. 711-734.

[17] Dorofeenko, V. G. Lee, and K. D. Salyer (2008): "Time-varying uncertainty and the credit channel", *Bulletin of Economic Research* 60(4), pp. 375-403.

[18] Furfine, C. H. (1999): "The microstructure of the federal funds market" *Financial Markets, Institutions & Instruments*, 8(5), pp. 24-44.

[19] Gale, D., and M. Hellwig (1985): "Incentive-compatible debt contracts: the one-period problem", Review of Economic Studies 52, pp. 647-664.

[20] Gambs, C. M., and D. V. Kimball: "Small banks and the federal funds market" *Economic Review*, Federal Reserve Bank of Kansas City.

[21] Goodfriend, M. and B. McCallum (2007): "Banking and interest rates in monetary policy analysis: A quantitative exploration" *Journal of Monetary Economics* 54, pp. 1480-1507.

[22] Goodhart, C. (2007): "Whatever became of the monetary aggregates?" LSE Financial Market Group Special Papers 172.

[23] Ho, T. S. Y., and A. Saunders (1985): "A micro model of the federal funds market" *Journal of Finance* 40(3), pp. 977-88.

[24] Jakab, Z. and M. Kumhof (2015): "Banks are not intermediaries of loanable funds - and why this matters" Bank of England Working Paper 529.

[25] Kiyotaki, N. and J. Moore (2002): "Evil is the root of all money" *American Economic Review* 92(2), pp. 62-66

[26] Holmes, A. R. (1972): "Operational constraints on the stabilization of money supply growth" in *Money and finance: readings in theory, policy and institutions*. Wiley, pp. 391-400.

[27] McLeay, M., A. Radia and R. Thomas (2014): "Money creation in the modern economy" Bank of England Quarterly Bulletin 54, First quarter.

[28] Sheard, P. (2013):"Repeat after me: Banks cannot and do not "lend out" reserves" Credit Market Services, Standard and Poor's, available at http://issuu.com/positivemoney/docs/sp-banks-cannot-anddo-not-lend-out.

[29] Stigum, M. and A. Crescenzi (2007): *Stigum's Money Market*, McGraw-Hill, 4th ed.

[30] Tobin, J. (1963): "Commercial banks as creators of 'money"', Cowles Foundation Discussion Paper 159, Yale University.

[31] Townsend, R. (1979): "Optimal contracts and competitive markets with costly state verification" *Journal of Economic Theory* 21, pp. 383-410.

[32] Werner. R. A. (2014): "How do banks create money, and why can other firms not do the same? An explanation for the coexistence of lending and deposit-taking" *International Review of Financial Analysis* 36, pp. 71–77.

[33] Werner. R. A. (2014): "Can banks individually create money out of nothing? The theories and the empirical evidence" *International Review of Financial Analysis* 36, pp. 1–19.

[34] Williamson, S. D. (1987): "Costly monitoring, loan contracts and equilibrium credit rationing" *Quarterly Journal of Economics* 102(1), pp. 135-146.

A Construction of $\Xi(x;\delta)$

Let the circle be divided in $2N$ identical contiguous intervals called villages so that each village covers a measure $1/(2N)$ of continuous locations. Let $x_v \in [0,1]$, be the fraction of entrepreneurs living in village $v \in \{1,2,\ldots,N\}$ who are able to pay back the bank loan. Assume all these fractions x_v, $v \in \{1,2,\ldots,2N\}$, are random variables with identical marginal distributions (not necessarily independent) denoted $\Xi(x)$. Let μ and σ be, respectively, the mean and standard deviation of $\Xi(x)$. Furthermore, assume the fractions x_v, $v \in \{1,2,\ldots,2N\}$, are correlated across villages in the following manner. The covariance between the fraction of solvent entrepreneurs in villages v and j, for all $v \in \{1,2,\ldots,2N\}$ and $j \in \{1,2,\ldots,2N\}$, is assumed to be

$$cov[x_v, x_j] = \sigma^2 \left[1 - 4\theta d(j,k)\right] \tag{48}$$

where $d(j,k)$ is the minimum distance, along the circumference of the circle, between villages v and j as measured by the distance between their center locations, and $\theta \in [0,1]$ is a parameter. Thus, the correlation of the fraction of solvent entrepreneurs gets smaller with distance with this correlation ranging from 1 (for arbitrarily close villages) to $1-2\theta$ (for locations opposite in the circle). The parameter θ controls the minimum correlation observed in the circle which happens between opposite locations.

Now take a general bank of measure δ meaning serving a number $2\delta N$ of contiguous villages. Let Δ be the set of contiguous villages this bank serves. Define the average solvency rate of entrepreneurs among *all villages* served by bank δ as

$$x(\delta) = \frac{1}{2\delta N} \sum_{v \in \Delta} x_v.$$

Because the realization of x_v is correlated among neighbor locations, this average default rate $x(\delta)$ is itself also a random variable following the probability distribution $\Xi(x;\delta)$. Let $\mu(\delta)$ and $\sigma(\delta)$ be, respectively, the mean and standard deviation of $\Xi(x;\delta)$. Clearly, the mean of this distribution is

$$\mu(\delta) = E\left[x(\delta)\right] = \frac{1}{2\delta N} \sum_{v \in \delta} E(x_v) = \mu$$

while, from (48), the variance is

$$\begin{aligned}
\sigma^2(\delta) &= Var\left[x(\delta)\right] = \frac{1}{(2\delta N)^2} \sum_{v \in \Delta} \sum_{j \in \Delta} cov[x_v, x_j] \\
&= \sigma^2 - \frac{4\theta\sigma^2}{(2\delta N)^2} \sum_{v \in \Delta} \sum_{j \in \Delta} d(v,j).
\end{aligned}$$

For $2\delta N \leq N+1$ (so that the bank could cover up to a little more than 50 percent of the loan market), it can be shown that

$$\sigma^2(\delta) = \sigma^2 - \frac{4\theta\sigma^2\delta}{(2\delta N)^2} \left[\frac{(2\delta N)^2 - 1}{3}\right].$$

Clearly, $\sigma^2(\delta)$ is decreasing in bank size, δ.[15] Also notice that, as $N \to \infty$, the variance settles at

$$\lim_{N\to\infty} \sigma^2(\delta) = \sigma^2 - \frac{4\theta\sigma^2\delta}{3}$$

which also decreases with the degree of diversification δ.

B The Loan Supply Schedule

Here I show how to derive the properties of the loan supply schedule from expressions (6)-(8) in the text. For convenience, these expressions are reproduced here:

$$\left(\frac{1+i^l}{1+i^o}\right)[\mu - F(\underline{x};\delta)] + \eta\left[(1+i^l)\left(F(\underline{x};\delta) - \phi G(\underline{x};\delta)\right) - 1 - i^o\right] = 0 \quad (49)$$

$$F'(\underline{x};\delta) = \eta(1+i^o)\left[F'(\underline{x};\delta) - \phi G'(\underline{x};\delta)\right], \quad (50)$$

and

$$(1+i^o)\lambda = \frac{1}{1 - \left(\frac{1+i^l}{1+i^o}\right)[F(\underline{x};\delta) - \phi G(\underline{x};\delta)]}. \quad (51)$$

The argument closely follows the derivations in Bernanke et al. [7]. First, notice that, for given rates i^o and i^l,

$$F'(\underline{x};\delta) - \phi G'(\underline{x};\delta) = 1 - \Xi(\underline{x};\delta) - \phi\underline{x}\xi(\underline{x};\delta) = [1 - \Xi(\underline{x};\delta)]\,[1 - \phi h(\underline{x};\delta)]$$

with

$$h(\underline{x};\delta) = \frac{\underline{x}\xi(\underline{x};\delta)}{1 - \Xi(\underline{x};\delta)}$$

being the hazard rate. By assumption i(iv), $h(\underline{x};\delta)$ is a strictly increasing function of $\underline{x}$. This means there must be a level of $\underline{x}$, call it $\underline{x}^*$, such that $F(\underline{x};\delta) - \phi G(\underline{x};\delta)$ is maximum which implies that leverage, λ, defined by (51) is also maximum. Thus, the relevant range to choose $\underline{x}$ from should be $\underline{x} \in [0, \underline{x}^*]$. No bank will choose a higher cutoff level $\underline{x}$ if, at the same leverage there is a lower level for $\underline{x}$ such as depositors are indifferent between the two. This is because the lower level for $\underline{x}$ saves on monitoring costs.

Use (50) to define the Lagrange multiplier η as a function of the cutoff level $\underline{x}$:

$$\eta(\underline{x};\delta) = \frac{F'(\underline{x};\delta)}{(1+i^o)\left[F'(\underline{x};\delta) - \phi G'(\underline{x};\delta)\right]}.$$

[15] The case $2\delta N > N + 1$ is a little bit more cumbersome to compute. In that case, however, the volatility of $x(\delta)$ is still decreasing in δ. As this case is not relevant empirically, it has not been worked out here.

Taking derivatives we have

$$\eta'(\underline{x}) = \frac{\phi\left[F'(\underline{x};\delta)G''(\underline{x};\delta) - F''(\underline{x};\delta)G'(\underline{x};\delta)\right]}{(1+i^o)\left[F'(\underline{x};\delta) - \phi G'(\underline{x};\delta)\right]} > 0.$$

for all $\underline{x} \in [0, \underline{x}^*]$. Furthermore, notice

$$\lim_{\underline{x}\to 0}\eta(\underline{x}) = \frac{1}{1+i^o};\ \lim_{\underline{x}\to \underline{x}^*}\eta(\underline{x}) = \infty.$$

Thus, expression (50) defines a one-to-one increasing mapping between the cutoff level $\underline{x}$ and the Lagrange multiplier η.

Next, from (49) define the function

$$1+i^l = \rho(\underline{x}) \equiv \frac{\eta(\underline{x})(1+i^o)}{\frac{\mu - F(\underline{x};\delta)}{1+i^o} + \eta(\underline{x})\left[F(\underline{x};\delta) - \mu G(\underline{x};\delta)\right]}.$$

Again, it can be shown that

$$\rho'(\underline{x}) \equiv \left[\frac{\rho(\underline{x})\eta'(\underline{x})}{(1+i^o)\eta(\underline{x})}\right] \frac{\mu - F(\underline{x};\delta)}{\frac{\mu - F(\underline{x};\delta)}{1+i^o} + \eta(\underline{x})\left[F(\underline{x};\delta) - \phi G(\underline{x};\delta)\right]} > 0$$

for all $\underline{x} \in [0, \underline{x}^*]$. Also, notice

$$\lim_{\underline{x}\to 0}\rho(\underline{x}) = \frac{1+i^o}{\mu};\ \lim_{\underline{x}\to \underline{x}^*}\rho(\underline{x};\delta) = \frac{1+i^o}{F(\underline{x};\delta) - \phi G(\underline{x};\delta)} < \frac{1+i^o}{\mu(1-\phi)}.$$

Thus, expression (49), together with (50), define a one-to-one increasing mapping between the cutoff level $\underline{x}$ and the loan rate i^l. Notice the previous expression bounds the loan rate between $(1+i^o)/\mu$ and $(1+i^o)/[\mu(1-\phi)]$.

Furthermore, use (51) to define the function

$$\lambda(\underline{x}) = \frac{1}{1+i^o} + \frac{\eta(\underline{x})\left[F(\underline{x};\delta) - \phi G(\underline{x};\delta)\right]}{\mu - F(\underline{x};\delta)}.$$

Taking derivatives one can be shown that

$$\lambda'(\underline{x}) \equiv \frac{\eta'(\underline{x})}{\eta(\underline{x})}\left[\lambda(\underline{x}) - \frac{1}{1+i^o}\right] + \frac{F(\underline{x};\delta)}{\mu - F(\underline{x};\delta)}\lambda(\underline{x}) > 0$$

for all $\underline{x} \in [0, \underline{x}^*]$ with

$$\lim_{\underline{x}\to 0}\lambda(\underline{x}) = \frac{1}{1+i^o};\ \lim_{\underline{x}\to \underline{x}^*}\lambda(\underline{x}) = \infty.$$

Thus, expression (51), together with (49) and (50), define a one-to-one increasing mapping between the cutoff level $\underline{x}$ and the leverage λ. Notice, as $1+i^l$ is reduced to $(1+i^o)/\mu$, the optimal cutoff level $\underline{x} \to 0$ and there will be no monitoring at the limit.

Finally, using (49) and (51), the excess return of bank equity over its opportunity cost,

$$\left(\frac{1+i^l}{1+i^o}\right)[\mu - F(\underline{x};\delta)]\,\lambda,$$

equals η.

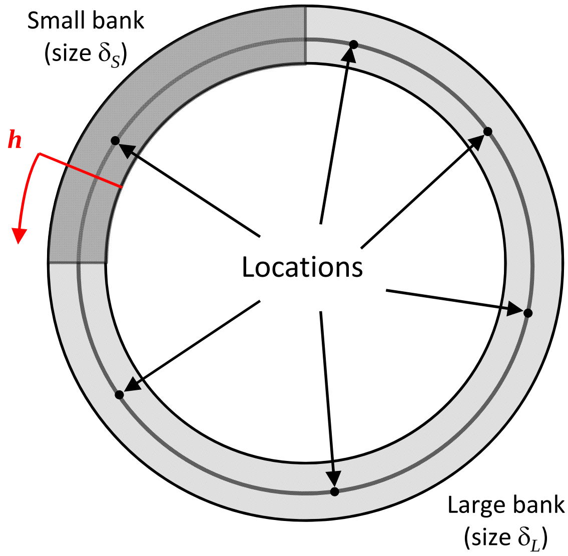

Figure 1: A representative island