Essays in Banking and Financial Structure

by

Swarnava Biswas

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Warwick Business School

July 2014

THE UNIVERSITY OF
WARWICK

Contents

Acknowledgments

As a new graduate student at Warwick Business School close to four years ago, I had very little idea of what to expect from my time here. Fortunately, it has been a very smooth ride. This has only been possible due to the help and support that I have received during my time as a PhD student.

First of all, I need to thank Kostas Koufopoulos for being a fantastic supervisor. It is fair to say that my entire research philosophy has been shaped by him. He has taught me the importance of identifying interesting questions and answering them rigourously. He is a famously harsh critic in general and is no different with me. And I am grateful for it as it has saved me a lot of time.

Andrea Gamba took over as my supervisor once Kostas left Warwick. Andrea has been very supportive and patient from the get go. Each time I went to him thinking I had something concrete, I left thoroughly confused from the barrage of question he would throw at me. Lucky for me he did that, as I wrote a paper (Chapter 3) trying to answer one of those many doubts that Andrea created in my mind.

I have benefitted greatly from discussions with Lei Mao who gave detailed comments on drafts. Anthony Neuberger has given me a lot of general advice over the years which holds me in good stead, going forward. Peter Corvi and Elizabeth Whalley have taught me a great deal about teaching. Jana Fidrmuc, Alessandro

Palandri and Vikas Raman have always had an open door so every time I have a doubt I have help at hand. Ingmar Nolte and Michael Moore have encouraged wholeheartedly to spend the department's money to go to conferences and I took full advantage. I believe that the exposure I gained eventually helped me get a job, so I am most grateful. I am also grateful to Sarojesh Mukherjee (my high school Economics teacher) who first ignited my interest in the subject.

My family has been a major pillar of strength throughout my life. The sacrifices that my parents have made to give me this opportunity to pursue a PhD is beyond what I could have imagined. They have supported my every endeavour, whether I succeeded or failed in the end. My brother, Sourav, has always given me the confidence to achieve more than I probably would have done otherwise. Brotivation aside, he also collected some data for me which although I haven't used in the thesis, has nevertheless been instrumental in helping me find a topic for a paper (Chapter 3). My fiancé, Emma, has been by my side through the process. I have bored her incessantly with my anecdotes from the academic world. I am just glad that she has decided to stick with me despite all that. To conclude, my thoughts go to Mickey, our family dog who passed away earlier this year at fourteen years of age. He was easily the most cherished member of the family and with good reason. I miss him dearly, now that he is gone.

Declarations

I declare that any material contained in this thesis has not been submitted for a degree to any other university.

I further declare that Chapters Two and Four of this thesis are a product of joint work with Dr. Kostas Koufopoulos.

Chapter Three is an independent paper. Having said that, the paper has benefitted immensely from numerous discussions I have had with Prof. Andrea Gamba, Dr. Kostas Koufopoulos and Dr. Lei Mao.

Swarnava Biswas

November, 2014

Abstract

The first essay (Chapter 2) highlights the positive effect of banks on direct financing in a setting where each agent believes that she can evaluate information better than any other agent. Banks emerge endogenously and they encourage direct financing through the use of underwriting and liquidity (reserve) requirements. Banks sell underwriting contracts to investors who wish to invest directly. Bank reserves reassures direct investors that the underwriting contract will be fulfilled. This results in the financing of positive NPV projects that were previously denied credit.

In the second essay (Chapter 3) an entrepreneur has the choice to access either monitored bank financing or un-monitored bond financing. Project type is private information of the entrepreneur and as a consequence, in the unregulated equilibrium, there is some inefficient over-monitoring by banks when the banking sector is competitive. Bank lending becomes more efficient and the net interest margin falls as bond financing becomes cheaper and the bond market expands. In contrast, if the banking sector is monopolistic, the equilibrium is either efficient or there is inefficient under-monitoring by banks.

The final essay (Chapter 4) proposes a model of optimal bank capital structure. There are two types of potential investors with different monitoring skills. The skilled can monitor a project and increase its productivity, whereas the unskilled cannot. Also, the skilled can divert a part of the project's return without being detected by the unskilled. Banks emerge endogenously and bank capital structure is relevant. The skilled become the bank's equity-holders whereas the unskilled become depositors. Our model explains why bank equity is more expensive than deposits.

Chapter 1

Overview

The thesis consists of three theoretical essays on the economics of banking and financial structure. Chapters Two and Three examine the choice of external financing source. Does the structure of the financial system - the mix of banks and financial markets - affect the real sector? While Chapter Four focuses on bank financing and explores security design issues. Across the thesis, the emphasis is to validate predictions using existing empirical results and generating new testable hypothesis.

The financial system comprises banks (more generally, financial intermediaries) and capital markets. The primary tasks of the financial system are related to resource allocation and risk-sharing. There is a large literature on the risk sharing aspect of the financial system (For example, Diamond and Dybvig, 1983, Allen and Gale, 1997).

In this thesis, I look at the role of the financial system in allocation of resources. Essentially, both banks and financial markets perform the same task which is transfer funds from one set of individuals (savers) to another (entrepreneurs with project ideas). Of course, the two mechanisms are distinct. While in bank financing, the investment decision is delegated to the bank manager, in financial markets the holders of funds make the decisions independently and invest directly.

One of the central issues in Chapters 2 and 3 is whether structure of the financial system (also known as the financial system architecture) matters for real growth. Does it matter whether an economy is market-based or bank-based? There is empirical evidence that not only the size, but also the structure of the financial system matters for economic growth (Tadesse, 2002). Further, Levine and Zervos (1998) find that both stock markets and banks positively affect economic growth.

At one end of the spectrum, the financial markets dominate the banking sector in countries like the US. At the other end, it is the banking sector which is dominant in countries like Japan and Germany. Is one system necessarily better than the other?

The second issue is whether there is a feedback channel between banks and financial markets. In most existing papers, banks and markets compete (For example, Allen and Gale (1997, 1999), Deidda and Fattouh (2008)). However, Demirguc-Kunt and Maksimovic (1996) and Demirguc-Kunt and Huizinga (2000) find cross-country empirical evidence that at least at the lower levels of financial development of an economy, banks and markets play a complementary role. In terms of theoretical work, Holmstrom and Tirole (1997) and Song and Thakor (2010) provide models of complementarity between banks and markets. In Holmstrom and Tirole (1997), the direction of feedback is one-way; bank financing allows firms to tap the capital market. In contrast, Song and Thakor (2010) provide a model in which the direction of feedback goes both ways.

Chapter 4 addresses a different question. It deals with security design issues in banking. Is bank equity more expensive than deposits (after risk adjustment)? In Chapters 2 and 3, the bank capital structure is irrelevant. Specifically, the bank may raise its funds from depositors or an equity base; no distinction is made between the two sources of funds. A standard assumption in banking literature is that bank equity is more expensive than deposits after risk adjustment. Recent examples include Allen, Carletti and Marquez (2011) and Mehran and Thakor (2011). The aim of Chapter 4 is to develop a framework in which bank equity capital is more expensive than deposits.

I discuss optimal intervention by a regulator whenever there is a market failure and an intervention may improve the net social welfare. In terms of government intervention, I take the stance that a regulator wishes to maximize the net social surplus. Other types of inefficiencies arise in the models (e.g. cross subsidization across types). However, the regulator does not care about the distributional inefficiencies. The sole objective of the regulator is to maximize the net social surplus.

Methodologically, the thesis uses simple game-theoretic models to study the role of banks. I use static games of external financing choice of firms to explore subtle and complex trade-offs and tease out important implications. The model in each of the three chapters has a distinct set of features and these are described below. However, there is a common thread which is present in all the three models in the thesis. The management of the bank is delegated to a manager who incurs a cost on behalf of the bank in order to resolve informational issues. The idea of delegation of monitoring has been around since Diamond (1984). As is well known, delegation is desirable on several accounts. In Chapters 2 and 3 it saves duplication of (learning or monitoring) costs and the free rider problem. In addition to cost-saving, in Chapter 4, the banker is more skilled or able to manage projects and increases the productivity of funds. Below, I outline the key features in the models and describe some results:

In Chapter Two (which is a coauthored project with Dr. Kostas Koufopoulos), delegation is costly because the uninformed bank members do not trust that the informed manager has accurately interpreted the data. Learning in the model is perfect, in the sense that if agents incur the learning cost they observe the future state with certainty and all informed individuals agree with one another. However, there is a behavioral assumption that uninformed individuals discount the decision-making skills of informed individuals. As a consequence, there may arise a state in which the informed banker observes the high state but fails to raise the necessary investment funds from the coalition members (who are uniformed). In such circumstances, the banker may sell underwriting (or insurance) contracts to capital market investors, as long as the banker can credibly promise that if the low state does occur, he will be able to fulfill the underwriting contract. Regulatory liquidity requirements make the underwriting contract credible and thereby, facilitate direct financing.

Our key contributions are as follows: banks emerge endogenously and they encourage direct financing through the use of underwriting and liquidity (reserve) requirements. This results in the financing of positive NPV projects that were previously denied credit. In addition, the paper provides a new rationale for liquidity requirements. The role of liquidity requirement in banking as a macro-prudential instrument is well understood (see Diamond and Dybvig, 1983). Our justification for the imposition of liquidity requirements comes from the asset side of the bank's balance sheet. We also highlight that there is a cost to imposing liquidity requirements as it makes the bank formation constraint tighter.

In Chapter Three, an entrepreneur may access either monitored bank financing or un-monitored bond market financing to raise funds for her project. While investors in the bond market only provide credit, a banker monitors a project in addition to provision of credit. Monitoring is costly but increases the probability with which a project succeeds. The key extension to the standard framework is that projects are heterogeneous; different projects experience different increase in success probability as a result of monitoring. It is efficient to monitor a project only if the net benefit from monitoring is positive. Project type is private information, which results in some inefficiencies in credit allocation.

In the unregulated competitive equilibrium, the banking sector is too large and there is over-monitoring. This inefficiency arises because some intermediate type projects (for which monitoring is socially wasteful) are subsidized in bank lending by higher type projects. I show that if bank equity is more expensive than deposits, a regulator may impose capital requirements (or tax the bank in another way) such that the first best outcome is achieved. On the other hand, if the banking sector is monopolistic, the equilibrium is either efficient (first best) or the banking sector is too small and there is under-monitoring. The key contribution here is that I relate lending efficiency to the net interest margin. I show that, irrespective of the structure of the banking sector, a wider net interest margin indicates higher degree of inefficiency in lending.

In Chapter Four (which is a co-authored project with Dr. Kostas Koufopoulos) we have two types of financiers with different monitoring skills. The skilled can monitor the project and increase its productivity, while the unskilled cannot. As a consequence, the skilled investors have a higher outside option than the unskilled, where the outside option is defined as the payoff an investor earns when investing on his own. Bank emerges as monitoring is delegated to a skilled investor due to economies of scale in monitoring. The key friction in the model is that a skilled banker may divert part of the profits from the unskilled bank members. The diversion is unverifiable in the court of law and may not be contracted upon.

In this risk-neutral setting bank capital structure is relevant and bank equity is more expensive than deposits. The first result is that the equity and deposit markets are segmented. While the skilled investors hold equity claim in the bank, the unskilled investors hold the senior deposit claim. If diversion is sufficiently large, an unskilled investor does not participate in the bank as an equity holder due to violation of his participation constraint. The senior debt claim gives the unskilled investors priority over the verifiable fraction of the cash-flow and convinces them to deposit in the bank. The expected return on equity is higher than the expected return on deposits. The higher return is not related to risk and arises due to scarcity of skill combined with a higher outside option of the skilled investors.

Bibliography

Allen, F. (1990): "The Market for Information and the Origin of Financial Intermediation," *Journal ofFinancial Intermediation*, 1, 3–30.

Allen, F., E. Carletti, and R. Marquez (2011): "Credit Market Competition and Capital Regulation," *Review of Financial Studies*, 24(4).

Allen, F., and D. Gale (1997): "Financial Markets, Intermediaries and Intertemporal Smoothing," *Journal of Political Economy*, 105(3), 523–546.

——— (1999): "Diversity of Opinion and Financing of New Technologies," *Journal of Financial Intermediation*, 8, 68–89.

Deidda, L., and B. Fattouh (2008): "Banks, Financial Markets and Growth," *Journal of Financial Intermediation*, 17(1), 6–36.

Demirguc-Kunt, A., and H. Huizinga (1999): "Determinants of Commercial Bank Interest Margins and Profitability: Some International Evidence," *World Bank Economic Review*, 13(2), 379–408.

Demirguc-Kunt, A., and V. Maksimovic (1995): "Stock Market Development and Financing Choices of Firms," *World Bank Economic Review*, 10(2).

Diamond, D. (1984): "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51, 393–414.

Diamond, D., and P. Dybvig (1983): "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, 91(3), 401–419.

Holmstrom, B., and J. Tirole (1993): “Market Liquidity and Performance Monitoring,” *Journal of Political Economy*, 101(4), 678–709.

Levine, R., and S. Zervos (1998): “Stock Markets, Banks and Economic Growth,” *American Economic Review*, 88(3).

Mehran, H., and A. Thakor (2011): “Bank Capital and Value in the Cross-Section,” *Review of Financial Studies*, 24(4).

Song, F., and A. Thakor (2010): “Financial System Architecture and the Co-evolution of Banks and Capital Markets,” *Economic Journal*, 120(547).

Tadesse, S. (2002): “Financial Architecture and Economic Performance: International evidence,” *Journal of Financial Intermediation*, 11(4), 429–454.

Chapter 2

The Beneficial Effect of Banks on Direct Financing

2.1 Introduction

The structure of the financial system is concerned with the channels through which funds are transferred from savers to entrepreneurs with investment opportunities. Some of the most pressing issues in this field relate to the determinants of the structure of the financial system - whether it is bank-based or market-based - and of course, the impact that this structure has on the real sector. There is evidence that not only the size, but also the structure of the financial system matters for economic growth (Tadesse, 2002). Further, Levine and Zervos (1998) find that both stock markets and banks positively affect economic growth. Therefore it is important to understand how the structure of the financial system mobilizes savings and channels them to positive NPV projects which, in turn, leads to higher economic growth.

In most existing papers, banks and markets compete.[1] In contrast, in our model banks have a positive effect on direct financing (markets). We identify two elements which are crucial for this positive effect: i) Underwriting or Insurance[2] and ii) Bank Liquidity or Reserve Requirements[3]. Banks encourage direct financing by selling underwriting contracts to investors who wish to invest directly. Bank reserves are set aside and not used for investment in projects. Instead they are used as a buffer stock to reassure agents who wish to invest directly that the underwriting contract will be fulfilled. We show that projects that were previously denied credit by either banks or directly, may receive funds with the help of underwriting and regulatory bank reserves. That is, this paper provides a new rationale for reserve/liquidity requirements. Our justification is based on the asset side of the bank's balance

[1] See for example, Allen and Gale (1997), (1999), Boot and Thakor (1997a), Deidda and Fattouh (2008)

[2] We use these terms interchangeably for the remainder of this text

[3] We use these terms interchangeably for the remainder of this text

sheet. This is in contrast to the existing literature where the arguments for reserve requirements come from the liability side.

More specifically, we consider a model where agents can pay a cost to learn the state of nature. If they observe the high state, they will invest in the project, whereas if they learn that the state of nature is low they will not invest. The formation of a coalition (bank) saves duplication of this learning cost as learning is delegated to a manager. Also agents underestimate the value of information when it is filtered through actions (underestimate the decision-making skills) of others. In any other respect, agents are rational and they maximize their utility given their perception of others' decision-making skills. If agents trust the decision-making skills of others sufficiently, they will coalesce to form a bank and delegate learning to a manager (a randomly picked individual with identical beliefs as the others in the population). Otherwise, each agent will incur the learning cost independently and subsequently decide whether or not to invest. Relative to direct financing, the trade-off for banks is lower learning cost (advantage) versus the possibility of disagreement about the project outcome (disadvantage).

Consider the case that prior to learning, it is optimal to form a bank. Ex-post however, coalition members may not sufficiently trust the manager and refuse to provide the required funds even if the manager has learnt that the state of nature is high. That is, because of the disagreement between the bank manager and the bank members, a project with positive net present value (NPV) does not receive financing although there are funds available. In this scenario, the manager can provide underwriting to investors who wish to invest directly. This insures these investors against the realization of the low outcome and incentivizes them to invest in projects that they would not otherwise.

What allows the bank manager to credibly offer underwriting is the regulatory bank reserves. In the absence of the bank reserves, the investors would not buy the insurance contract, as according to their perception it could not be fulfilled. Thus, the use of underwriting backed by the bank reserves implies that more positive NPV projects receive direct financing. This leads to a larger financial system and higher economic growth (because more positive NPV projects receive financing). This is consistent with existing empirical evidence (e.g. Levine, 2001). However, despite their usefulness, bank reserves are not provided voluntarily by the bank members and they should be imposed by regulation.

Imposing liquidity requirements on a bank is not without a cost. We show that higher liquidity requirements make the bank formation constraint tighter and so, there is an upper bound to liquidity requirements consistent with bank formation. With a perfect learning technology the banks reserves are never used; it is simply there to make the banker's promise credible. However, we show that when the learning technology is noisy, bank reserves are used on the equilibrium path to cover unexpected losses (see Section 4.2).

The key to our results is the assumption that individuals undervalue the information of others. This assumption may not be consistent with rational expectations and so it requires some justification. Numerous experimental studies, both by psychologists and economists, indicate that the majority of people overestimate their abilities and the outcomes of their actions.

In one of the earliest studies, Svenson (1981) finds that 93% of the automobile drivers in Sweden consider themselves 'above average'. Camerer and Lovallo (1999) show in an entry game that subjects overestimate their relative ability to solve puzzles. Malmendier and Tate (2005, 2008) provide evidence that CEOs are overconfident about their abilities to manage a company. These and related studies are summarized in DellaVigna (2009).

Moreover, Weizsacker (2003) finds that agents significantly underestimate the rationality of others. In effect, people tend to be overly skeptical of the decision-making skills of others. Also, Weizsacker (2010) finds that in situations where it is optimal to follow others and contradict one's own information, the players still err in the majority of cases, forgoing substantial parts of earnings (rejecting rational expectations). Importantly, Andreoni and Mylovanov (2012) find that disagreements may persist over time even when sufficient information to reach agreement is provided. This occurs because agents put more weight on the information that they receive directly, rather than indirectly through actions of others.[4]

In our setting, all agents who incur the learning cost reach the same conclusions and never disagree with one another. However, agents are skeptical about the conclusions of others (from learning) and believe that they may have drawn different (more accurate) conclusions, from identical information if they were to observe the information themselves. We further assume (in line with Allen and Gale, 1999), that the degree of skepticism is project-specific. The clarity of the data will determine the level of skepticism about the decision-making skills of others. An innovative project with complex and scarce data will be more difficult to value and individuals will trust less the decision-making skills of others regarding this project.

[4]Santos-Pinto and Sobel (2005) provide a theoretical explanation to this behavior.

Regarding empirical implications, our model predicts that projects characterized by a lot of uncertainty (and hence disagreement) will be directly financed while those with low uncertainty are financed by banks. We share this prediction with Allen and Gale (1999). Further, our model predicts that a decrease in the degree of uncertainty about the prospects of projects leads to a higher level of bank liquidity requirements consistent with bank formation.

2.2 Related Literature

Our paper is related to three strands in the literature. i) The literature on reserve requirements, ii) The literature on financial intermediation and iii) The literature on the interaction between banks and capital markets.

i) Central banks explicitly state that they impose reserve requirements to foster financial stability and to deal with volatile capital flows (Glocker and Towbin, 2012). Calomiris (2011) argues that it is important to think of reserve requirements as more than simply an instrument to limit illiquidity risk. Calomiris, Heider and Hoerova (2012) propose a model in which liquidity requirements are used as a macro-prudential instrument. In the interbank market, banks commit to lending each other funds to insure against bank-specific liquidity needs. Liquidity requirements are imposed to prevent free-riding on the efficient interbank liquidity assistance. Additionally, higher cash reserves are imposed to mitigate risk-shifting incentives. Moreno (2011) and Gray (2011) describe the use of reserve requirements as a macro prudential tool. In contrast, our arguments for liquidity requirements stem from the asset side of the bank's balance sheet. Liquidity requirements allows the bank to credibly sell underwriting contracts (expand its assets) which result in the financing of positive NPV projects, which are otherwise denied credit.

ii) There is a large literature on the existence of financial intermediaries. The existence of financial intermediaries is usually attributed to their informational advantages; they facilitate pre-contract information production and post-contract monitoring in a way that lenders avoid duplication costs and entrepreneurs exert the desired effort level. Another strand of the literature provides liquidity provision stories for the existence of financial intermediaries.[5] The papers more closely related to ours are the following:

Our intermediary is similar in spirit to the one in Diamond (1984). The intermediary in our model avoids duplication of the learning cost (pre-lending as opposed to post-lending moral hazard resolution in Diamond, 1984) by delegating the task to a manager. However, there is also a cost to intermediated finance which arises due to agents' skepticism over the decision-making skills of others. The trade-off determines which form of financing is more efficient.

Allen (1993) notes that banks and markets perform two unique and separate roles in the economy, thereby emphasizing the importance of each. What distinguishes banks and markets is the structure of institutions as opposed to the instruments they use. Banks individually negotiate contracts while in capital markets there are large numbers of anonymous lenders who take the pre-specified contracts. Markets allocate resources efficiently - provided pricing is accurate - because they continuously check whether the manager is working sensibly and that he or she has information

[5]See for example, Leland and Pyle (1977), Diamond and Dybvig (1983), Diamond (1984), Ramakrishnan and Thakor (1984), Boyd and Prescott (1986), Allen (1990), Gorton and Penacchi (1990), Allen and Gale (1997), and Diamond and Rajan (2000,2001) among others

close to the true information set or at least the market's perception of the true information set. If a manager is far off from the market's beliefs, prices will reflect that as a lot of different views - those of the market participants - will come together as opposed to a single bank doing the monitoring. In that sense, capital markets provide a way of checking whether firms are well-run when there are diverse opinions on the project's prospects. When there is no consensus on how firms should be run, banks may not be as effective as there is a single check which is performed by the bank's manager. Hence banks are a good way to provide financing in traditional industries where technology is well known and there is a wide consensus on how firms should be run. In this case, the bank can monitor firms effectively and take advantage of scale economies in monitoring. Financing through capital markets is thus desirable in industries where there is a high degree of technological innovation. This idea is formalized in Allen and Gale (1999) which forms the basis of our model. Our paper differs from Allen and Gale (1999) in three key aspects: First, our interpretation of the disagreement among agents is different.[6] Second, in Allen and Gale (1999) the agents know the exact value of β (degree of agreement) before they decide whether they invest on their own or coalesce to form a bank. In contrast, in the generalized version of our model (section 3), when the agents make this decision they know the distribution of β but not the exact value. They observe the exact value of β only after becoming a bank member. Third, Allen and Gale (1999) do not consider bank reserves and underwriting and so, in their model there is no interaction between banks and markets.

iii) Boot and Thakor (1997a,b) provide the first formal model that endogenizes the co-existence of banks and markets by highlighting the uniqueness of each. Borrowers

[6]See the discussion in the introduction (page 5) and in the model.

with lower observable quality prefer bank financing as banks mitigate the post-lending asset substitution problem and these borrowers do not suffer a loss with banks. Borrowers who pose milder moral hazard risk will go directly to capital markets as informed investors are reasonably sure that these borrowers will invest in good projects and make use of their information (which is conveyed through prices). In these models, banks and markets compete which is in contrast with our paper, in which banks and markets positively interact.

Song and Thakor (2010) develop a model in which banks and markets complement each other and co-evolve. Through the securitization channel, banks increase participation in capital markets. As participation in the capital market increases, the bank's cost of raising equity capital falls which leads to an expansion of bank lending. In our case, banks expand direct financing through underwriting contracts which are backed by the regulatory bank reserves.

2.3 The Model

2.3.1 The Economic environment

Our model is very similar to Allen and Gale (1999). More specifically, we consider a three-date economy ($t = 0, 1, 2$) with universal risk neutrality and a zero riskless interest rate. There are two types of agents: entrepreneurs and investors; all agents consume at $t = 2$. The entrepreneur has access to a project that needs investment, I at $t = 1$ and returns are realized at $t = 2$. The project yields H in the good state and L in the poor state ($H > I > L$). The probability of each state is specific to each project and not a function of the overall economy. The entrepreneur has zero endowment and seeks the required funds, I, for investing in the project at $t = 1$. There are X investors who are symmetric in beliefs regarding the profitability of the project, to start with. Each of these investors has an endowment of $(I + 2c + K)$ at $t = 0$. At date $t = 0$, the investor has the choice to incur a fixed cost $c > 0$, to obtain further information prior to deciding whether or not to invest. A project has constant returns to scale and its scale is capped. The maximum investment in one project is YI. We assume that the sum of the endowments of all investors is less than the amount demanded by entrepreneurs. This assumption implies that the financiers (individual investor or the bank) have all the bargaining power and so the full surplus (the NPV of the project) accrues to the financiers.

There are also some other agents that have zero endowment and the bank manager will be chosen among these agents.[7] Managerial compensation is a positive function of the ex-post bank profit and is paid at $t = 2$. Compensation is an infinitesimally small proportion of the realized returns (manager's outside option is normalized to 0) and we generally ignore it in our analysis as we do not have any implications for/from it.

By incurring the fixed cost of learning, c at $t = 0$ any individual will learn the state of nature prior to deciding whether or not to invest, I at $t = 1$. If she observes the high state of nature, the project returns H at $t = 2$, otherwise the project returns L. Learning is therefore perfect in this model. The ex-ante probability of the high state is given by α.

[7] Alternatively, the bank manager could have been randomly chosen among the bank members. The only difference is that in case of disagreement, the bank manager could invest his own funds, I. Assuming that the maximum scale of the project, YI, is sufficiently larger than I, investment I is negligible. However, our current formulation simplifies the algebra.

To make the analysis interesting, we make the following assumption:

Assumption 1: $\alpha H + (1-\alpha)L - I < 0$ i.e. it is never optimal to invest unconditionally.

This assumption implies that no project will receive any funding, unless either individual investors or a bank incurs the cost of learning the state of nature.

If all agents incur the learning cost, c, they will all draw the same conclusion regarding the outcome of the project. However, agents discount information when it is filtered through the actions of others. An uninformed agent believes that the conclusion reached by another individual is correct with probability $\beta \in [\alpha, 1)$ and wrong with the complementary probability. These beliefs are common knowledge among all agents in the economy. Furthermore, we assume that the parameter, β is project specific. An innovative project with complex and scarce data will be more difficult to value and individuals will trust less the decision-making skills of others regarding this project (lower β). If $\beta = \alpha$, the agents do not learn anything from the actions of others and do not update their unconditional probability despite observing the actions of others who have analyzed the data. Any project is characterized by the two parameters, (α, β). Although our model is similar to Allen and Gale (1999), there are two critical differences with respect to the parameter β. First, our interpretation of β is different. More specifically, Allen and Gale define it as the measure of correlation among informed investors' beliefs. That is, in their interpretation, if two agents incur the cost and learn the state of nature, they may still disagree with probability, $(1-\beta)$, whereas in our interpretation they fully agree. In our case, the disagreement arises only among an informed and an uninformed agent. Second, in the generalized version of our model (section 3.3), when the agents decide whether to invest on their own or coalesce to form a bank, they do not know the exact value of β, as in Allen and Gale (1999). They only know the distribution of β. They observe the exact value of β at $t = 1$. The idea behind this more general formulation is that between bank formation and the actual investment in the project some more information arrives which helps the uninformed bank members obtain a better estimate of their degree of disagreement with the bank manager. Allen and Gale (1999) implicitly assume that no new relevant information becomes available to bank members after bank formation.

2.3.2 Basic model

For expositional purposes we start with the benchmark case where the exact value of the project β is known at $t = 0$, there are no liquidity requirements and no underwriting.

Direct Financing

Direct financing occurs when investors incur the cost of learning the state of nature and then decide individually whether or not to fund the project. Because $H > I > L$, an individual invests only if she observes the high state in which the project returns H. The expected payoff at $t = 0$ to an informed individual investor if she invests directly is,

$$V_D = \alpha(H - I) - c \tag{2.1}$$

If an investor decides to invest c at $t = 0$, at $t = 1$ she learns the state of nature. Therefore at $t = 0$, an individual investor incurs the learning cost only if,

$$V_D = \alpha(H - I) - c > 0$$

Bank Financing

Intermediated financing is the case when N investors will coalesce to share the learning cost, c; each member of the coalition will contribute $\frac{c}{N}$ towards the learning cost at $t = 0$ and if at $t = 1$ perceives it as a profitable investment, she will provide I units of investment funds. An investor has the right to participate in the investment of the project through the bank only if he pays his share of the learning cost at $t = 0$. Information gathering is delegated to a manager who is randomly selected from the population. The manager will become informed and truthfully reveal his findings to the members at $t = 1$, who then decide whether to provide the investment funds or not. The coalition members do not learn the state of nature themselves - the manager can only reveal to them his own opinion about the prospects of the project. With probability β, the coalition members believe that the manager has assessed the project accurately and with probability $(1 - \beta)$ coalition members believe that manager has made a mistake. If the manager observes the high state, the expected return of the investment for an uninformed investor is $\beta H + (1 - \beta)L$.
Hence, the expected payoff at $t = 0$, to a bank member is,

$$V_B = \alpha(\beta H + (1 - \beta)L - I) - \frac{c}{N} \tag{2.2}$$

At $t = 0$, coalition formation is feasible as long as the expected return is weakly positive. So, each member will provide her share of the learning cost, $\frac{c}{N}$ only if,

$$V_B = \alpha(\beta H + (1-\beta)L - I) - \frac{c}{N} > 0$$

If this condition is satisfied and the manager observes the high state after learning, all coalition members provide the investment amount, I, at $t = 1$ because the above condition ($V_B > 0$) implies that,

$$\beta H + (1-\beta)L - I > 0$$

The learning cost is treated as sunk at $t = 1$. Note that if the opinion of others is not undervalued i.e. ($\beta = 1$), intermediated finance strictly dominates direct financing, as duplication of learning costs is costlessly avoided (the bank disadvantage disappears).

Direct vs Bank Financing

At $t = 0$ individual investors decide whether they will incur the learning cost, c, and invest on their own or not invest at all or they will coalesce with other investors and share the learning cost. If some agents decide to coalesce but the expected payoff of each member is lower than what she can achieve on her own (because the number of agents is not sufficiently high) then the coalition is dissolved and each individual acts on her own. We look for symmetric pure strategy Nash Equilibria in this game. The agents choose their strategies to maximize their expected payoffs and this determines the form of financing.

Depending on parameter values, there are four cases to be considered:

Case 1: $V_B < 0$ and $V_D < 0$ for any $N \leq Y$.

Case 2: $V_D > 0$ and $V_D > V_B$ for any $N \leq Y$.

Case 3: $V_B > V_D > 0$ for some $N \leq Y$.

Case 4: $V_B > 0 > V_D$ for some $N \leq Y$.

In cases 1 and 2 the outcome is clear as it is a strictly dominant strategy for all agents. In case 1, the unique equilibrium involves no financing. In case 2, there is a unique equilibrium where only direct financing exists.

In cases 3 and 4 the outcome is described in the following proposition:

Proposition 1: *Let $V_B > 0 > V_D$ or $V_B > V_D > 0$ for some $N \leq Y$. Then the efficient outcome (bank financing) arises as a unique Nash equilibrium. Only one bank finances each project (The bank size, NI, equals the size of the project YI) and each individual member incurs a learning cost of $\frac{c}{y}$.*

Proof. Given that $V_B > 0 > V_D$ or $V_B > V_D > 0$ for some $N \leq Y$, the expected payoff for an individual from joining a coalition is $pV_B + (1-p)V_D > V_D$ or $pV_B > 0$ for any $0 < p \leq 1$, where p is the probability that the investment takes place through the coalition (bank). So, for any agent, it is a strictly dominant strategy to join some other agent(s) and form a coalition. We now show that there cannot exist an equilibrium in which the bank size (NI) is less than the scale of the project (YI). Consider the case where there are two banks of equal size ($\frac{YI}{2}$). In this case, a member of either bank can increase her profit by joining the other bank because now she shares the cost, c, with more members. By a similar argument, we can rule out any bank size but, YI. Suppose now that there is one bank with size $NI = YI$.

Then the deviation of any member of this bank to either market financing or no financing would imply a lower profit for her (given the conditions in cases 3 and 4). Therefore, the equilibrium with bank financing with one bank is unique. □

2.3.3 A more general model

In this section we extend the basic model in two dimensions:
i) The project's β is not known with certainty at $t = 0$. We assume that β is uniformly distributed in the interval $[\underline{\beta}, \overline{\beta}]$ and this is public knowledge at $t = 0$. The exact value of β is observed at $t = 1$.
ii) The manager may undertake off-balance sheet activities such as underwriting

The payoff of direct financing is exactly the same as in the previous section. For simplicity , we consider the case that the expected payoff to direct financing is negative ($V_D < 0$). This implies that direct financing is infeasible and so the outside option of an investor is who does not join a bank is 0. We consider bank financing for the following two cases:
i) There are no liquidity requirements and
ii) Regulators impose a liquidity requirement (defined later)

Direct Financing

The payoff of direct financing in the market is as before. The expected payoff to an informed individual investor in the market is,

$$V_D = \alpha(H - I) - c$$

We assume that $V_D < 0$. That is, direct financing in infeasible. In what follows, we show that the existence of banks can prompt direct financing in an environment where the latter does not otherwise occur (although it is efficient).

Bank Financing with No liquidity requirements:

First, we consider the case in which there are no regulatory liquidity requirements. At $t = 0$, it is commonly known that the project's β is uniformly distributed in $[\underline{\beta}, \overline{\beta}]$, and the bank members observe the exact value of β at $t = 1$ (denote as β_R).

Definition: There is a threshold, β^*, where $\underline{\beta} < \beta^* < \overline{\beta}$, for which the net payoff to providing investment funds to the manager at $t = 1$ is 0 (after treating learning costs as sunk) and is given by,

$$\beta^* H + (1 - \beta^*)L - I = 0 \Rightarrow \beta^* = \frac{I - L}{H - L} \tag{2.3}$$

At $t = 0$, the bank members provide the learning cost only if,

$$V_B = \alpha \left[\frac{\beta^* + \overline{\beta}}{2} H + \left(1 - \frac{\beta^* + \overline{\beta}}{2} \right) L - I \right] - \frac{c}{N} > 0 \tag{2.4}$$

Consider the case that the manager observes the high state. If $\beta_R > \beta^*$, the coalition members provide the required funds, I at $t = 1$ since,

$$\beta_R H + (1 - \beta_R)L - I > 0$$

If, however, $\beta_R < \beta^*$, the coalition members do not provide the required funds at

$t = 1$ since,

$$\beta_R H + (1 - \beta_R)L - I < 0 \tag{2.5}$$

Thus, at $t = 1$, if $\beta_R < \beta^*$, the manager cannot invest in the project, despite observing the high state. The manager wishes to maximize the ex-post profit for the bank to maximize his own compensation. To that end, given that he observes the outcome as high with certainty, he attempts to sell underwriting contracts to the direct investors.

Direct investors have the same information as the bank members; i.e. they observe β_R. If we assume that the direct investors do not observe β_R, our arguments still go through. As we describe below, the key to our argument is that direct investors have a sufficiently low β (i.e. $\beta < \beta^*$). Even if they did not observe the exact realization of β, direct investors observe that the bank members do not provide the investment funds and rationally infer that $\beta_R < \beta^*$. Their best conditional estimate of the project β is $\frac{\beta+\beta^*}{2} < \beta^*$.

Definition: An underwriting contract is a commitment at $t = 1$ of the bank to the direct investor that in the event that the low state, L occurs at $t = 2$, the bank will transfer some funds to the investor, such that the return to the investor in the low state is $L' > L$. To buy this contract, the investor pays a fee, Z, to the bank at $t = 1$. The expected payoff of the individual investor in the market from a project financed through underwriting is denoted V_C,

$$V_C = \beta_R H + (1 - \beta_R)L' - I - Z \tag{2.6}$$

Assuming that the bank has all the bargaining power, the net payoff to the market investors will be 0 ($V_C = 0$). This implies that the fee is given by,

$$Z = \beta_R H + (1 - \beta_R)L' - I \tag{2.7}$$

In order to maximize the bank profit (and his compensation), the bank manager will maximize Z given $L' \leq H$. [8]

Lemma 1: *The underwriting contract sets $L' = H$ and so the fee, $Z = H - I$.*

Proof. See Appendix. □

Market investors will accept the underwriting contract only if the bank manager has access to some funds, which will reassure them that the manager will fulfil his promise. However, given $\beta_R < \beta^*$ the bank members will not voluntarily provide any funds more than c at $t = 0$ and no market investor will buy the underwriting contract. Formally,

Lemma 2: *At $t = 0$, the bank members will never voluntarily provide more than the learning cost, c to the bank manager.*

Proof. See Appendix. □

Lemma 3: *In the absence of bank liquidity requirements, direct investors will not buy the underwriting contract offered by the bank manager.*

Proof. See Appendix. □

[8] Notice that the underwriting contract may be purchased even by the bank members who have refused to provide funds directly to the bank.

Bank Financing with Regulatory liquidity requirement

From above, it is clear that the bank manager will have, at $t = 0$, in his availability, funds more than c, only if it is imposed exogenously by a regulatory authority.

Regulator: Consider now a regulator whose objective is to maximize the net social surplus. The regulator can take the following actions: First, he can learn the state of nature by incurring the learning cost c at $t = 0$. This cost is covered by the bank members. The regulator incurs the cost, c and therefore, she observes the state of nature at t=1. As a consequence, at t=1, the regulator perfectly agrees with the bank manager regarding whether to invest in the project or not. In effect, the regulator's beta is 1[9]. Second, he can impose liquidity requirements, K, per member on the intermediary (as described below). The regulator, however, does not have the managerial skill to run a bank.

Regulatory liquidity requirements: The regulator (government) requires that in order to form a coalition, investors have to provide an additional amount K, on top of the total learning cost, $2c$ (learning cost is $2c$ since both the banker and the regulator incur this cost). The bank reserves will be kept on the bank's balance sheet and the manager cannot use it for direct investment. Liquidity requirements have been put in place by the government to buffer the bank against unexpected losses and allow the manager to carry out off-balance sheet activities, such as underwriting. At $t = 2$ bank members get their funds back in full, as long as no losses have been incurred on off-balance sheet activities.

[9]This is a simplifying assumption and the results are robust even if regulator learning is not perfect. The results hold as long as the regulator's β is greater than β^*. It is natural to assume that the regulator's β is greater than the β of the bank members, since the regulator incurs the learning cost and learns the state of nature, where as the bank members do not.

If $\beta_R > \beta^*$, the ex-ante expected payoff is similar to that in equation (2.4), with the only difference being that instead of c we have $2c$ (the additional c is the learning cost of the regulator),

$$\alpha\left[\frac{\beta^*+\overline{\beta}}{2}H+\left(1-\frac{\beta^*+\overline{\beta}}{2}\right)L-I\right]-\frac{2c}{N} \tag{2.8}$$

Let us consider the case where $\beta_R < \beta^*$. Suppose that the manager observes the high state. The coalition members do not sufficiently trust the manager and do not provide the investment funds, I. The manager offers the underwriting contract to market investors. With regulatory reserves in place, the manager's promise is credible, since he has access to the bank reserves and if the bad state occurs, he can resort to these reserves to make the promised payments. Thus, direct investors buy some underwriting. The coalition members take this into account at $t = 0$ when they provide the learning cost ($t = 0$) and the regulatory reserves to form the bank. The condition for bank formation becomes stricter as investors explicitly take into account the ex-ante (perceived) losses that they incur if underwriting contracts are sold (of course, ex-post the underwriting contract is in fact profitable for coalition members as well).
If $\beta_R < \beta^*$, the coalition members do not provide the investment funds and the manager extends an underwriting contract using the regulatory reserves, K, as a buffer.

From Lemma 1, the underwriting contract specifies that the buyer receives H regardless of the state realized. Hence, the manager needs to have access to funds worth $(H - L)$, so she can reassure the buyer that if the low outcome realizes the buyer will be compensated according to the terms of the underwriting contract.

Each bank member provides an additional amount K (on top of his share of the learning costs). Thus, if $K \geq H - L$, the bank manager has enough funds available to credibly sell the maximum number of underwriting contracts, Y, (which equals the maximum scale of the project and the bank size). If $K < H - L$, the bank manager can at most sell $\frac{K}{H-L}$ underwriting contracts per bank member or $\frac{K}{H-L}Y$ underwriting contracts in total (which is less than the maximum scale of the project). With probability $\left(\frac{\underline{\beta}+\beta^*}{2}\right)$, the bank members agree with the manager and expect the high state to occur. There are no payouts to the buyers of the underwriting contract if the high state occurs. Of course, the bank earns the fee from the underwriting contract, $(H - I)$ (from Lemma 2). However, according to the perceptions of the bank members, with probability $\left(1 - \frac{\underline{\beta}+\beta^*}{2}\right)$, there is a loss of $(H - L)$ per unit of underwriting contract sold. Combining with the fee (which is paid irrespectively of the state), the net payoff is $(L - I)$. Therefore, the (perceived) expected payoff to each bank member in this case is given by,

$$\alpha \left[min \left(\frac{K}{H-L}, 1 \right) \left[\left(\frac{\underline{\beta} + \beta^*}{2} \right) (H - I) + \left(1 - \frac{\underline{\beta} + \beta^*}{2} \right) (L - I) \right] \right] - \frac{2c}{N} \quad (2.9)$$

From the definition of β^*, this perceived payoff is negative. Combining equations (2.8) and (2.9), we obtain the perceived expected payoff for each member at $t = 0$.

$$V_B = \alpha \left[\frac{\beta^* + \overline{\beta}}{2} H + \left(1 - \frac{\beta^* + \overline{\beta}}{2} \right) L - I \right] +$$
$$\alpha \left[min \left(\frac{K}{H-L}, 1 \right) \left[\left(\frac{\underline{\beta} + \beta^*}{2} \right) (H - I) + \left(1 - \frac{\underline{\beta} + \beta^*}{2} \right) (L - I) \right] \right] - \frac{2c}{N} \quad (2.10)$$

Because the expression in equation (2.9) is negative, the higher the regulatory liquidity requirements, K, the lower the perceived expected payoff to bank members. So, the necessary condition for bank formation ($V_B \geq 0$) becomes tighter. This condition gives us an upper bound on K consistent with bank formation:

$$V_B \geq 0 \Rightarrow K \leq -\frac{\alpha\left[\frac{\beta^*+\overline{\beta}}{2}H + \left(1-\frac{\beta^*+\overline{\beta}}{2}\right)L - I\right] - \frac{2c}{N}}{\frac{\alpha}{H-L}\left[\left(\frac{\underline{\beta}+\beta^*}{2}\right)(H-I) + \left(1-\frac{\underline{\beta}+\beta^*}{2}\right)(L-I)\right]} \equiv \overline{K} \tag{2.11}$$

Since the regulator also observes the state of nature, he always agrees with the bank manager. As a result, in order to achieve his objective, the regulator chooses the reserve requirement K so as to maximize the amount of investment in the positive NPV projects, given that the necessary condition for bank formation is satisfied. Investors may still want to join the bank at $t = 0$ because the expected payoff of investing in the project through the bank is positive. The proposition below summarizes these results:

Proposition 2: *If $\overline{K} \geq H - L$, then the regulator imposes $K \geq (H - L)$. If $\overline{K} < H - L$, the regulatory reserves are set at $K = \overline{K}$.*

Notice that if $K > H - L$, the manager only uses $(H - L)$. The remaining amount $(K - H + L)$ is not used. Hence, if the condition for bank formation ($V_B \geq 0$) is satisfied for $K \geq H - L$, an imposition of a higher K does not make this condition tighter.

Recall that in the absence of bank liquidity requirements, the bank members do not provide any funds beyond $\frac{c}{N}$ to the bank. If at $t = 1$, the bank manager observes the high state but $\beta_R < \beta^*$ the coalition members do not provide the investment funds, I. The bank manager cannot sell underwriting contracts because market investors do not perceive her promise as credible. Therefore, a positive NPV project is not financed. With bank liquidity requirements, the bank manager sells some underwriting contracts (since market investors are reassured by the bank reserves) and the buyers of these contracts invest in the project. More specifically,

Proposition 3: *If $\overline{K} \geq H - L$, the bank manager sells Y underwriting contracts and the investment in the project is YI (its maximum scale). If $\overline{K} < H - L$, the bank manager sells $\frac{\overline{K}}{H-L}Y$ underwriting contracts and the investment in the project is $\frac{\overline{K}}{H-L}YI$ (less than its maximum scale).*

Proof. See Appendix. □

Intuitively, the bank reserves allow the bank manager (through the underwriting contracts) to credibly promise the market investors the high outcome, regardless of the realization of the state. Thus, market financing occurs even though in the absence of banks it would not be possible (since $V_D < 0$). By reducing the cost of learning (per capita) from c to $\frac{2c}{Y}$, the bank converts a negative NPV project into a positive NPV one. However, the disagreement between the bank manager and bank members does not allow the bank to invest directly in the project. The bank reserves imposed by the regulator allows the bank manager to sell underwriting contracts which incentivize market investors to invest in a project that was previously denied credit from both the market and the bank. That is,

Corollary: *The combination of economies of scale in learning, regulatory bank reserves and underwriting contracts allows the bank manager to trigger direct financing in a project that would not otherwise be funded.*

The key to our model is the assumption that individuals undervalue information when it is filtered through actions of others, i.e. $\beta < 1$. The underwriting contract is a costless exercise for the manager who observes the outcome is high while it is valuable to the market investor; trade is mutually beneficial due to this difference in perceptions.

Further, the underwriting contract can only be credibly extended if there are liquidity requirements in place. A benevolent regulator is aware that liquidity requirements are overall welfare improving. Thus, liquidity requirements (imposed by a regulator) endogenously arise in this model. To reiterate a point made earlier, the regulatory reserve requirement is not without a cost - it makes the condition for bank formation stricter.

We should also point out that although banks can trigger direct financing, the opposite is not true in the model. The number of projects funded directly and the profit investors make on these projects have no effect on the bank-formation constraint (equation 2.4). This constraint is only affected by the degree of (dis)agreement among the agents in the economy, β, which is project-specific and the ex-ante probability of the high state, α. Thus, the profit the agents make by directly investing in projects cannot trigger bank formation.

We can derive an interesting and potentially testable empirical predictions by comparative statics with respect to the expected β.

Lemma 4: *Consider the case in which, $\overline{K} < H - L$. An upward shift in the distribution of β (keeping the difference between the two bounds constant), leads to a higher $\overline{K}$.*

Proof. See Appendix. □

That is, the lower the uncertainty about the project's prospects, the higher the bank liquidity requirements consistent with bank formation.

2.4 Discussion

2.4.1 Credit Default Swaps

Here we show that the insurance provided by the bank reserves cannot be obtained through other means (cannot be created endogenously and has to be imposed). Consider, for example, the case that there are no regulatory reserves in the bank's balance sheet and that the bank buys insurance through Credit Default Swaps (CDS).

The CDS issuer optimally incurs the learning cost, c, and so he agrees with the bank manager regarding the outcome of the project. However, the direct investor who buys the underwriting contract from the bank, agrees with the bank manager and the CDS issuer with a probability $\beta_R < \beta^*$. So, the investor will be convinced to buy the underwriting contract only if the CDS issuer has set aside the funds required for fulfilling his promise. But, in raising these funds, the CDS issuer faces exactly the same problems as the bank manager. The perceived payoff of the fund providers is given by equation (2.9), which is clearly negative. As a result, they do not supply the required funds to the CDS issuer and the CDS contract is not credible to the buyers of the underwriting contract. Hence the reserves required to credibly sell the underwriting contract must be imposed by regulation.

2.4.2 Noisy Learning

Thus far, we have considered the case that the signal is perfect and if one incurs the learning cost, s/he perfectly foresees the future state of nature. As a consequence, the bank never pays out on the underwriting contract and the bank reserves are never used. Suppose now, that the signal is noisy: in the high state, payoff is H with probability p and L with probability $(1 - p)$. In this case, the banker is wrong with probability $(1 - p)$ and the bank reserves are used to pay the buyers of the underwriting contracts with the same probability. The implication is that the condition for bank formation becomes stricter. However, qualitatively, our results remain unchanged.

2.5 Conclusion

We analyze a simple model of a financial system in which all agents are rational but they discount information when it is filtered through actions of others. In our model banks emerge endogenously and positively affect direct investment. This interaction is facilitated by the use of underwriting and regulatory bank reserves. Bank reserves are used as a buffer stock to reassure direct investors that the underwriting contract will be fulfilled. This leads to an increase in direct financing - more positive NPV projects are undertaken. However, the bank reserves are not provided by bank members voluntarily and have to be imposed by regulation. In the absence of bank liquidity requirements, underwriting would be infeasible as the bank manager could not have made credible promises regarding payments if the poor outcome occurred. Thus, our model provides a new rationale for bank liquidity regulation. The novelty is that our justification is based on the asset side of the bank's balance sheet as opposed to the liability side in the existing literature. Further, our model predicts that a decrease in the degree of uncertainty about the prospects of projects, an upward shift in β, leads to a higher level of bank liquidity requirements consistent with bank formation.

Finally, although in this paper we have considered a static model, our basic idea can be embedded into a dynamic context to obtain a two-way relationship between banks and direct financing (markets). In the dynamic setting, direct financing can have a positive effect on banks as the profits made by banks on the underwriting contracts can be carried forward to relax their future funding constraint. Also, additional concerns such as the banker's reputation and the bank's payout policy will become important. We leave these issues for future research.

2.6 Appendix: Proofs

Proof of Lemma 1: The underwriting contract stipulates that the transfer $(L'-L)$, will be made in the event of poor outcome. The manager knows with certainty that the high outcome will be realized and so there are no future payments related to the underwriting contract. Therefore, to maximize Z, the manager will fully insure the buyer against the poor outcome; i.e. $L' = H$. Using $L' = H$ in equation (2.6), we obtain $Z = H - I$.

Proof of Lemma 2: If at $t = 0$, any funds in excess of the learning cost, c is provided, at $t = 1$ a manager who has observed the high state, will use the excess funds to invest in the project (either directly or through underwriting) even if $\beta_R < \beta^*$. But the bank members perceive this strategy as loss-making. Furthermore, they can always provide the funds after the realization of β. That is, for bank members, any funds in excess of c available to the bank manager imply a perceived cost (when $\beta_R < \beta^*$) without any compensating benefit. Therefore, at $t = 0$, they will not voluntarily provide any funds more than c.

Proof of Lemma 3: By Lemma 2, the bank manager has no funds available after incurring the learning cost, c. Market investors believe that the bad state may occur with probability $(1 - \beta_R)$ and in that event the bank manager will not be able to fulfil his promise. As a result for $\beta_R < \beta^*$, their perceived payoff is $\beta_R H + (1 - \beta_R)L - I - Z < 0$ for any $Z \geq 0$ as $\beta_R H + (1 - \beta_R)L - I < 0$.

Proof of Proposition 3: In equilibrium, Y agents coalesce to form the bank ($N = Y$). By Proposition 2, if $\overline{K} \geq H - L$, the regulator imposes $K \geq H - L$. Hence, in this case, the bank manager can credibly promise to fulfill Y underwriting contracts ($H - L$ is the amount that the bank has to pay the buyer of the underwriting contract when the bad state realizes) and so, the project is undertaken at its maximum scale. If $\overline{K} < H - L$, the regulator sets $K = \overline{K}$ (by Proposition 3). As a result, the bank manager can credibly promise to fulfil at most $\frac{\overline{K}}{H-L}Y$ underwriting contracts which implies that the investment in the project is $\frac{\overline{K}}{H-L}YI$.

Proof of Lemma 4: The maximum K per bank member consistent with bank formation is given in equation (2.11). An increase in $\overline{\beta}$ increases the numerator whereas an increase in $\underline{\beta}$ makes the denominator less negative (the expected perceived loss decreases). As a result, $\overline{K}$ increases.

Bibliography

Allen, F. (1990): "The Market for Information and the Origin of Financial Intermediation," *Journal ofFinancial Intermediation*, 1, 3–30.

——— (1993): "Stock Markets and Resource Allocation," in *Capital Markets and Financial Intermediation*, ed. by C.Mayer, and X.Vives, Cambridge, UK. Cambridge Univ. Press.

Allen, F., and D. Gale (1997): "Financial Markets, Intermediaries and Intertemporal Smoothing," *Journal of Political Economy*, 105(3), 523–546.

——— (1999): "Diversity of Opinion and Financing of New Technologies," *Journal of Financial Intermediation*, 8, 68–89.

Andreoni, J., and T. Mylovanov (2012): "Diverging Opinions," *American Economic Journal - Microeconomics*, 4(1), 209–232.

Boot, A., and A. Thakor (1997a): "Financial System Architecture," *Review of Financial Studies*, 10(3), 693–733.

——— (1997b): "Banking Scope and Financial Innovation," *Review of Financial Studies*, 10, 1099–1131.

Boyd, J., and E. Prescott (1986): "Financial Intermediary Coalitions," *Journal of Economic Theory*, 38(2), 211–232.

Calomiris, C. (2011): "Identifying the right mix of Capital and Cash Requirements in Preudential Bank Regulation," *Working Paper*.

Calomiris, C., F. Heider, and M. Hoerova (2012): "A Model of Bank Liquidity Requirements," *Working Paper*.

Camerer, C., and D. Lovallo (1999): "Overconfidence and Excess Entry: An Experimental Approach," *American Economic Review*, 89(1), 306–318.

Coval, J., and A. Thakor (2005): "Financial Intermediation as a Beliefs-bridge between Optimists and Pessimists," *Journal of Financial Economics*, 75, 535–569.

Deidda, L., and B. Fattouh (2008): "Banks, Financial Markets and Growth," *Journal of Financial Intermediation*, 17(1), 6–36.

DellaVigna, S. (2009): "Psychology and Economics: Evidence from the field," *Journal of Economic Literature*, 47(2), 315–372.

Diamond, D. (1984): "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51, 393–414.

——— (1997): "Liquidity, Banks and Markets," *Journal of Political Economy*, 105, 928–956.

Diamond, D., and P. Dybvig (1983): "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, 91(3), 401–419.

Diamond, D., and R. Rajan (2000): "A Theory of Bank Capital," *Journal of Finance*, 55, 2431–2465.

——— (2001): "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking," *Journal of Political Economy*, 109, 287–327.

Glocker, C., and P. Towbin (2010): "Reserve Requirement for Price and Financial Stability: Why are they effective?," *International Journal of Central Banking*, 8(1), 65–113.

Gorton, G., and G. Penacchi (1990): "Financial Intermediaries and Liquidity Creation," *Journal of Finance*, 45(1), 49–71.

Gray, S. (2011): "Central Bank balances and Reserve Requirements," *IMF Working Paper*.

Leland, H. E., and D. H. Pyle (1977): "Information Asymmetries, Financial Structure and Financial Intermediation," *Journal of Finance*, 32, 371–387.

Levine, R. (2001): "Bank-based or Market-based Financial Systems: which is better?," *Journal of Financial Intermediation*, 11(4), 398–428.

Levine, R., and S. Zervos (1998): "Stock Markets, Banks and Economic Growth," *American Economic Review*, 88(3).

Malmendier, U., and G. Tate (2005): "CEO Overconfidence and Corporate Investment," *Journal of Finance*, 60(6), 2661–2700.

——— (2008): "Who makes Acquisitions? CEO Overconfidence and the Market's Reaction," *Journal of Financial Economics*, 89(1), 20–43.

Moreno, R. (2011): *Policymaking from a "macroprudential" perspective in emerging market economies*, BIS working papers. Bank for International Settlements, Monetary and Economic Dept.

Ramakrishnan, R., and A. Thakor (1984): "Information Reliability and a theory of Financial Intermediation," *Review of Economic Studies*, 51(3), 415–432.

Santos-Pinto, L., and J. Sobel (2005): "A Model of Positive Self-Image in Subjective Assessments," *American Economic Review*, 95(5).

Song, F., and A. Thakor (2010): "Financial System Architecture and the Coevolution of Banks and Capital Markets," *Economic Journal*, 120(547).

Svenson, O. (1981): "Are we all less risky and more skillful than our fellow drivers?," *Acta Pschycologica*, 47, 143–147.

Tadesse, S. (2002): "Financial Architecture and Economic Performance: International evidence," *Journal of Financial Intermediation*, 11(4), 429–454.

Thakor, A. (1996): "The Design of Financial Systems: an overview," *Journal of Banking and Finance*, 20(5), 917–948.

Weizacker, G. (2003): "Ignoring the Rationality of Others: Evidence from Experimental Normal-form games," *Games and Economic Behaviour*, 44, 145–171.

——— (2010): "Do We Follow Others when We Should? A simple Test of Rational Expectations," *American Economic Review*, 100(5).

Chapter 3

Credit Market Efficiency and the Net Interest Margin

3.1 Introduction

Consider a situation in which a firm (the borrower) has the choice between direct (bond) and indirect (bank) financing. In this paper, the goal is to examine the impact of the firm's external financing choice on the overall efficiency in the credit allocation (lending) process. I propose that a single measure, the bank net interest margin, captures the degree of efficiency in the overall lending environment, in a setting in which borrower type is unobservable. Further, I relate the structure of the banking sector to issues of efficiency.

Motivation: Bernanke (1983) notes that the cost of credit intermediation (which could be the physical cost of screening or monitoring) in the banking sector reflects the degree of efficiency in bank lending. A higher cost leads to a credit squeeze by the banking sector which adversely affects the real sector.

In empirical work, Demirguc-Kunt and Huizinga, (1999, 2000) proxy bank lending efficiency using net interest margins. In their words, the bank net interest margin reflects the spread between the net returns to savings and the gross return for real investment. The argument goes that a wider net interest margin indicates a higher cost in transferring funds from savers to borrowers and hence (cost) inefficiency.

I develop a model in which net interest margins reflect not only cost efficiency in bank lending but also informational efficiency in the overall credit allocation process. In a sense, I relate the bank net interest margin to the net social welfare.

Summary of the Model: I consider a model in which an entrepreneur (she) seeks external financing but cannot signal her type, which leads to inefficiency. The entrepreneur may approach either a bank or the bond market to raise funds for the project.

In bond financing, the investor simply provides the requisite funds to undertake the project. With bank financing, in addition to providing credit, the banker (he) also provides monitoring services. This may be due to a free-rider problem in the bond market which is only solved in a coalition. Alternatively, banks could have access to proprietary information which the bond market investors do not (Bhattacharya and Chiesa, 1995; Yosha, 1995). Finally, it could be because they do not possess the monitoring technology or that monitoring is prohibitively costly for bond market investors.

Any project, if not monitored, succeeds with probability, p, and fails with probability, $(1-p)$. Monitoring a project increases the probability with which it succeeds by δ. However, monitoring is costly and the cost is incurred upfront (at date, 0). Heterogeneity among the projects arises as follows: δ differs across projects and some projects experience a higher increase in probability of success due to monitoring than others. It is efficient to monitor a project only if the net benefit from monitoring is positive.[1] δ represents the private information of the borrower. See the Literature Review section for a detailed motivation of the assumptions regarding the parameter, δ.

In a competitive lending environment, the financiers compete for projects and earn a zero expected profit while the borrowers keep the entire surplus. The projects for which monitoring is most efficient, borrow from the bank as they keep the surplus from monitoring. The projects for which monitoring is least efficient, borrow in the bond market.

[1]The assumption is formally stated in $A2$, page 10.

Some intermediate type projects may find it profitable to pool with the higher types to take advantage of better borrowing terms. It is privately profitable for these projects to demand bank loans, even though monitoring is not efficient for them from a social perspective. This occurs when the benefit from a lower repayment in bank financing exceeds the expected loss from monitoring, for a project.

I show that the unregulated competitive equilibrium is *always* inefficient and is characterized by some degree of over-monitoring. Furthermore, a smaller net interest margin corresponds to the equilibrium moving towards the efficient outcome.

I show that if bank equity is more expensive than deposits, regulatory bank capital (up to a certain level) will increase the net social welfare. Expensive equity imposes a cost on bank financing, which makes it less attractive. There is an optimal level of bank capital which may be imposed to achieve the efficient outcome. If a higher requirement is imposed, the equilibrium will be inefficient and characterized by under-monitoring. Critically, the bank equity has to be imposed as banks will not optimally hold any costly equity.

Finally, I consider a monopolistic banking sector. In contrast to above, I show that the monopolistic equilibrium may in fact be efficient. It is also possible that the banking sector is inefficiently small (characterized by under-monitoring by banks as opposed to over-monitoring which is the case with a fully competitive banking sector). However, as before, a smaller interest margin indicates a move towards the more efficient outcome. The source of inefficiency in this case in monopoly pricing and informational asymmetry does not play a role.

The contributions of the paper are as follows:

1. If the banking sector is perfectly competitive, there is over-monitoring. At the other end of the spectrum, if the banking sector is monopolistic, there may be the socially optimal level of monitoring or too little monitoring.

2. I show that a single measure, the bank net interest margin, captures the degree of efficiency in the overall credit allocation process. Specifically, a wider bank net interest margin indicates greater social inefficiency, all else equal. Further, this result holds irrespective of whether the banking sector is perfectly competitive or monopolistic (although, the mechanisms are distinct).

3. I provide a rationale for imposing costly bank capital requirements when the banking sector is competitive. Costly equity makes bank financing less attractive and pushes the equilibrium towards the efficient outcome. There is an optimal level of equity that is imposed by the regulator to achieve the efficient outcome.

3.2 Related Literature

De Meza and Webb (1987) consider a similar mechanism in the context of a competitive banking sector. Their results critically rely on the specific assumptions on the supply of funds to the banking sector (the supply of deposits). If the supply of deposits is not decreasing in the rate of return on deposits, they show that banks over-invest. If, however, the supply of deposits is backward-bending, banks under-invest in equilibrium. Theoretically, the key difference in this paper is that I explicitly model the deposit market which allows me to explore the net interest margin.

The present paper is related most closely to two separate strands of the literature: **i.** the role of banks and **ii.** the choice between informed and arm's length debt.

i. Role of banks

The key assumption in the model may be broken down into two separate components as follows: bank monitoring is valuable (there exists a positive δ) and different projects benefit to different degrees from bank financing (δ differs across projects).

Positive δ. Fama (1985) notes that there must be something special about bank loans that borrowers are prepared to bear the costs of reserve requirements that do not appear in direct financing. He suggests that contracting costs for bank loans may be lower than for market securities (like bonds). Further, bank loans may carry a signalling effect which improves creditworthiness and reduces information costs of other contracts.

Chemmanur and Fulghieri (1994) present a model in which banks endogenously produce more information than the bond market to build reputation. More information allows banks to make the liquidation decision more efficiently. By doing so, they build reputation over time. Bond market investors do not have these reputation building incentives as they are single period players.

Ramakrishnan and Thakor (1984) and Allen (1990) emphasize the role of banks as information producers for outside agents (such as depositors). Leland and Pyle (1977), Diamond (1984) and Boyd and Prescott (1986) show that banks provide more effective and efficient monitoring services.

Empirically, James (1987) finds that bank loan announcements by firms are associated with positive market reactions, while announcements of market debt are associated with zero or negative market reactions. More recently, Altman, Gande and Saunders (2010) and Gande and Saunders (2012) find similar results. Gande and Saunders (2012) argue that in addition to resolution of information asymmetries a borrowing firm benefits from a banking relationship when the bank sells its loans in the secondary loan market.[2] Banks add to firm value through loan sales in the secondary market which alleviates the borrower's financial constraints.

In sum, Boot (2000) notes that banks develop close relationships with borrowers over time and this proximity facilitates monitoring and screening to overcome costs associated with asymmetric information.

Heterogeneity in δ. The feature in my model which drives the core results, is that δ differs across projects. In a sense, I relax the standard assumption that all projects benefit equally from bank monitoring. I do not explicitly model the source of heterogeneity.

Empirically, Gande and Saunders (2012) find that smaller and financially distressed borrowers have a higher positive stock reaction when its loans are sold in the secondary market by their bank. In terms of the model presented here, it may be interpreted that these borrower have a higher δ and benefit more from bank financing.

I have assumed that the parameter, δ, is unobservable; i.e., it is the private information of the borrower. While there are observable proxies for financial constraints, these proxies are noisy. I have assumed un-observability to simplify the analysis. The core results go through as long as δ is not perfectly observed.

[2] Although, I call it bank monitoring, I am silent on the specific channels through which bank adds value to a borrower. The value-added from bank financing may come from access to the secondary loan market; however, I do not explicitly model this feature.

ii. The choice between informed and arm's length debt:

Besanko and Kanatas (1993) present a model in which bank lending and other credit contracts co-exist. Like I have in the model here, in addition to providing credit a bank adds value through monitoring; while direct financing only provides credit. Since the bank cannot commit to a level of monitoring, firms are financed with a mixture of bank loan and direct financing (bonds or equity).

Diamond (1991) looks at the role of the firm's reputation in its choice of external financing source. Firms with high credit ratings rely on their reputation to raise funds in the bond market. While, firms with middle of the spectrum credit ratings borrow from the bank which incurs costly information acquisition and gives the firm a better deal than uninformed bond market investors. Extremely low rated firms rely on bond market financing (junk bonds).

Chemmanur and Fulghieri (1994) look at the reputational concerns of banks and show that firms which are at a higher risk of facing temporary negative shocks choose bank financing since they have better incentives to produce costly information and make the right liquidation decision.

Rajan (1992) suggests that bank's private information lets them hold up borrowers in the refinancing stage and extract rents. A firm trades off the usual benefits of bank lending (such as resolution of information asymmetries) against the cost of hold up which determines its choice of financing source.

Holmstrom and Tirole (1997) and Repullo and Suarez (2000) study how the firm's net worth relates to its choice between direct and indirect financing. Morellec, Valta and Zhdanov (2013) and von Thadden (1994) examine how the choice of external financing source affect firm's investment behavior.

Boot and Thakor (1997a,b) provide a model of coexistence of banks and equity markets. Song and Thakor (2010) show that evolution of the equity markets make (expensive) regulatory capital cheaper for banks which allows them to extend credit to riskier borrowers.

These papers generally do not address security design issues (neither do I). One exception is Bolton and Freixas (1999) in which bank loans, bond financing and equity financing co-exist in equilibrium.

3.3 Model

3.3.1 Set-up

I consider a one-period ($t = 0, 1$) economy in which all agents are risk-neutral. All returns are consumed at the end of the period. The risk-free rate is normalized to 0, so there is no discounting. In the core model, there are four types of agents: the borrower/entrepreneur, the bank, depositors and bond market investors.

The entrepreneur (she) has access to a project. She is penniless and seeks outside financing for her project. The project requires an investment, I, which is normalized to 1. The investment is undertaken at $t = 0$ and returns to the project are realized at $t = 1$.

The entrepreneur may approach either a bank or the bond market to raise funds for the project. The key difference between the two types of financing is that while investors in the bond market only provide credit, a banker monitors the project in addition to providing the credit. It is assumed that in either case, the full investment capital, 1, is borrowed. Further, investment is verifiable in the court of law and everyone who raises finance, invests.

A project returns X (success) or 0 (failure). If a project is not monitored (bond financing), it succeeds with probability, p and fails with probability, $(1-p)$.

I consider only positive NPV projects:

A1: $pX - 1 > 0$

If, on the other hand, a project is monitored (bank financing), the probability with which it succeeds increases by δ (payoffs in the different states are unaffected); the project succeeds with probability, $(p+\delta)$ and fails with probability, $(1-p-\delta)$.

δ differs across projects and is the solitary source of heterogeneity among projects within the model. There are some projects which are uniformly distributed in the interval, $\delta_i \in [0, \bar{\delta}]$. $\bar{\delta}$ satisfies the restriction, $(p+\bar{\delta}) \leq 1$. The assumption of uniform distribution is for algebraic simplicity and does not carry qualitative implications.

The increase in success probability of the project due to monitoring, δ_i, is private information of the entrepreneur. All projects are ex-ante observationally equivalent and there are no screening technologies to ascertain a project's type.

Monitoring is costly. To monitor a project, a non-pecuniary cost, M, is incurred by the bank at date 0. The NPV of monitoring a project of type i is $\delta_i X - M$. There exists a project, δ_m, for which the NPV of monitoring is equal to 0.

A2: $\delta_m X - M = 0$ for some $\delta_m \in [0, \bar{\delta})$

There is a cut-off δ_m such that monitoring is only efficient if monitoring increases the success probability of the project by δ_m (or more) and it is inefficient to do so otherwise.

The assumption states that there are some projects, $\delta_m < \delta_i \leq \bar{\delta}$ for which the bank adds a strictly positive value. If the monitoring cost, M, is high such that $\delta_m \geq \bar{\delta}$, bank monitoring does not add a strictly positive value for any project. I do not consider this case.

For now, it is assumed that the monitoring cost is observable and verifiable by all. If this were not the case, monitoring can be made incentive compatible by requiring the banker to invest some of his own wealth in the project (as in Holmstrom and Tirole, 1997). I show this later (Section 3.7).

A3: Projects are scarce.
I consider a perfectly competitive credit environment where financiers (banks or bond market investors) compete for the projects. Financiers earn zero profit in expectation. The entrepreneur chooses the mode of financing to maximize her payoff. Effectively, all costs are internalized by the entrepreneur.

In a variant of the core model, I consider the case in which the banker behaves monopolistically and extracts some rent. (Section 5)

3.3.2 Bond Financing

The entrepreneur may raise the investment capital, 1, in the bond market (think of the bond market investors as a single party). The investor offers a contract, $(R_c, 0)$

to ensure that she breaks even. R_c is the required repayment if the project succeeds and 0 if the project fails.[3] In the downside, the entrepreneur is protected by limited liability.

The bond market investor does not monitor the project and therefore, it succeeds with probability, p. In a competitive credit market, an investor in the bond market earns a zero profit in expectation,

$$pR_c - 1 = 0 \tag{3.1}$$

From the zero profit condition, we obtain the required repayment in bond financing, R_c:

$$R_c = \frac{1}{p} \tag{3.2}$$

The borrower's expected payoff, denoted π_c, is the probability of success, p times the net project payoff which is X minus the scheduled repayment R_c, i.e.,

$$\pi_c = p(X - R_c) = pX - 1 \tag{3.3}$$

Given $A1$ (all projects are positive NPV), an entrepreneur's expected payoff to bond financing is always positive.

[3] The subscript c indicates Capital market.

3.3.3 Bank Financing

Suppose that the average project seeking bank loan is of type, δ_b (δ_b will be determined in equilibrium).

Deposits

The banker raises the investment funds in the deposit market. Supply of deposit is infinite and the depositor earns a zero profit. Each depositor deposits an amount, 1.

The zero profit condition of the depositor is given as follows (where R_D is the deposit rate):

$$(p + \delta_b)R_D - 1 = 0 \tag{3.4}$$

From the zero profit condition of the depositor, we derive the deposit rate,

$$R_D = \frac{1}{p + \delta_b} \tag{3.5}$$

Loans

The bank offers a contract, $(R_L, 0)$ to the borrower to ensure that it breaks even. R_L is the required repayment if the project succeeds and 0 if the project fails. In the downside the entrepreneur is protected by limited liability.

The banker incurs the monitoring cost which is M. In a competitive financing environment, the banker earns a zero profit in expectation,

$$(p + \delta_b)(R_L - R_D) - M = 0 \tag{3.6}$$

Substituting R_D in the zero profit condition, we obtain the loan rate:

$$R_L = \frac{1 + M}{p + \delta_b} \tag{3.7}$$

The borrower's expected payoff, at $t = 1$ is the probability of success, $(p + \delta_b)$ times the net project payoff which is X minus the scheduled repayment R_L. The net payoff, denoted π_i, is given as follows:

$$\pi_i = (p + \delta_i)(X - R_L) \tag{3.8}$$

Given that the borrower's expected payoff in bond financing is always positive, in equilibrium the same is true for bank financing. This is true because a project with negative expected payoff in bank financing will simply raise funds in the bond market instead.

3.3.4 Relative Cost of Debt

Bank debt is more expensive than bond financing if,

$$R_L(\delta_b) - R_c > 0 \Rightarrow M > \frac{\delta_b}{p} \tag{3.9}$$

If the monitoring cost is sufficiently large, bank debt is more expensive (consistent with previous literature, e.g., Holmstrom and Tirole, 1997).

Additionally, bank financing adds (weakly) positive value, $\delta_i X$, to the project. Therefore, if equation (3.9) is not satisfied, there is no bond financing in equilibrium and all projects are financed by the banking sector.

Equation (3.9) is a necessary but not sufficient condition for bond financing.

3.3.5 The Efficient Equilibrium

In this section, I define the efficient equilibrium. I show in the rest of the paper that due to asymmetric information regarding borrower type, the competitive equilibrium is never efficient in an unregulated economy, although the efficient outcome may arise in the monopolistic banking sector.

I define the efficient equilibrium as the one in which there is no over or under monitoring. In the efficient equilibrium all projects for which monitoring is efficient, are monitored and projects for which monitoring is inefficient, are not monitored. Specifically,

Definition: The efficient equilibrium is one in which any project with $\delta_i < \delta_m$ seeks financing in the bond market and any project with $\delta_i > \delta_m$ seeks a bank loan. The case of a project with $\delta_i = \delta_m$ is inconsequential to efficiency considerations.

In Chemmanur and Fulghieri (1994), more bank loans invariably leads to higher social efficiency since banks acquire more information and liquidate projects more efficiently. The difference in this model is that while it is efficient for some projects to go to banks, it is efficient for the other projects to go to the bond market.

There is another source of inefficiency in the model: in bank financing, the more profitable projects subsidize the less profitable projects (higher δ projects subsidize the lower δ projects), as they accept a higher repayment schedule relative to the full information case (where the contract is written on observable type). However, this inefficiency is related to distribution of surplus and does not affect the net social surplus. I disregard it in further analysis.

3.3.6 Equilibrium

I assume that the borrowers and the financiers play the following two-stage game:

Stage 1: The banker and the bond market investors simultaneously offer a contract, $(R_j, 0)$ where $j = c, L$.
Stage 2: Given the offers made by the financiers, the entrepreneurs apply for either bank or bond financing.

Definition: We consider a pure-strategy sub-game perfect equilibria. An equilibrium consists of a choice of the borrowing contract by the entrepreneur and the choices of interest rates by the banks and bondholders, which satisfies the following conditions: given equilibrium choices and beliefs of the other players, **a.** an entrepreneur's choice maximizes her expected payoffs and **b.** the interest rates set by the lenders (depositors lending to banks and banks and bondholders to the entrepreneurs) are the lowest that gives them non-negative expected profits in equilibrium. Any deviation from equilibrium strategies by any player results in a lower expected payoff for him/her compared to that obtained in equilibrium, given others' beliefs.

Proposition 1: *There exists a cutoff,* $\delta_i = \delta_q$*, where,*

$$\delta_q = max\left[0, \frac{1}{2X}\left[-((2p+\bar{\delta})X-(2M+1)) \pm \sqrt{((2p+\bar{\delta})X-(2M+1))^2 + 4X(2pM-\bar{\delta})}\right]\right] \tag{3.10}$$

An entrepreneur will only seek bank loan if $\delta_i > \delta_q$ *and prefer bond financing if* $\delta_i < \delta_q$.[4] *She is indifferent if* $\delta_i = \delta_q$.

Proof. Assume initial beliefs, δ_q. An entrepreneur seeks bank financing if $\delta_i > \delta_q$ and bond financing if $\delta_i < \delta_q$. She is indifferent if $\delta_i = \delta_q$.

The average project borrowing from the bank succeeds with a probability, $(p + \delta_b(\delta_q))$. Therefore, in Stage 1 of the game the bank sets the borrower's repayment (from equation (3.7)) as,

$$R_L(\delta_q) = \frac{(1+M)}{p+\delta_b} \text{ where } \delta_b = \frac{\delta_q + \bar{\delta}}{2} \tag{3.11}$$

An entrepreneur, with a project of type, δ_i, is indifferent between the two sources of financing if her expected payoffs are identical in both cases, for the given beliefs:

$$pX - 1 = (p+\delta_i)[X - R_L(\delta_q)] \tag{3.12}$$

The LHS is the expected payoff for an entrepreneur who borrows in the bond market, irrespective of her type. The RHS is the expected payoff for an entrepreneur who borrows from the bank and the payoff is increasing in type, δ_i.

It must be true that the above equation only holds for the marginal project,

[4] If both roots of δ_q are positive, either one may be the equilibrium.

$\delta_i = \delta_q$. From the above, we derive an expression for the cutoff, δ_q (equation (3.10)).

Note that the equation is quadratic in δ_q. Restrictions on M are imposed to ensure that a positive root exists. The details of the derivation are put in the Appendix.

By construction, equation (3.12) holds with equality for $\delta_i = \delta_q$. For any $\delta_i > \delta_q$, the RHS is greater than the LHS; the entrepreneur strictly prefers bank loan to bond financing. And finally, for any $\delta_i < \delta_q$, the LHS is greater than the RHS; the entrepreneur strictly prefers bond financing to bank loan.

Therefore, there is no profitable deviation for any project from the above equilibrium and the initial beliefs, δ_q, are proved to be correct. □

In an efficient equilibrium, δ_q and δ_m coincide. However, we see below that such an equilibrium does not exist in an unregulated economy and that the equilibrium will be characterized by some degree of inefficiency.

Proposition 2: *The equilibrium is inefficient and is characterized by over-monitoring by banks, i.e., $\delta_q < \delta_m$.*

Proof. The proof is by contradiction.

Substituting $R_L(\delta_q)$ and R_c in equation (3.12), the condition for indifference is rewritten as,

$$pX - 1 = (p + \delta_i)\left[X - \frac{2(1+M)}{2p + \delta_q + \bar{\delta}}\right] \tag{3.13}$$

Suppose that $\delta_q = \delta_m$.

Consider the marginally efficient project, $\delta_i = \delta_m$. Substituting and rearranging equation (3.13),

$$\frac{(\delta_m - \bar{\delta})}{2p + \delta_m + \bar{\bar{\delta}}} = \delta_m X - \frac{2(p + \delta_m)M}{2p + \delta_m + \bar{\bar{\delta}}} \tag{3.14}$$

Note that $\delta_m X = M$ (by Assumption $A2$). Substituting,

$$\frac{(\delta_m - \bar{\delta})}{2p + \delta_m + \bar{\bar{\delta}}} = M - \frac{2(p + \delta_m)M}{2p + \delta_m + \bar{\bar{\delta}}} \tag{3.15}$$

Note that $LHS < 0$, since $\delta_m < \bar{\delta}$. However, $RHS > 0$ since $\frac{2(p+\delta)}{2p+\delta_m+\bar{\delta}} \leq 1$ for any $\delta_m < \bar{\delta}$.

Therefore, for $\delta_i = \delta_m$, equation (3.15) is violated. A project with $\delta_i = \delta_m$, strictly prefers to borrow from the bank. But, this is a contradiction to the starting point, $\delta_q = \delta_m$.

Similarly, it can be shown that $\delta_q \not> \delta_m$. It follows that $\delta_q < \delta_m$. □

The high δ_i ($\delta_i > \delta_m$) projects seek bank financing. The low δ_i ($\delta_i < \delta_q$) projects seek bond financing.

The most interesting case is that of the intermediate project ($\delta_q < \delta_i < \delta_m$):

It is socially wasteful to monitor these projects, i.e., $\delta_i X - M < 0$. However, pooling with the higher types subsidizes the intermediate type projects. The subsidy makes bank financing privately optimal for these projects, i.e., $\delta_i X - (R_L(\delta_q) - R_c) \geq 0$; with equality for $\delta_i = \delta_q$.

3.3.7 Net Interest Margin

In this section, I derive the net interest margin in bank lending and discuss some key empirical implications.

Definition: The net interest margin is the difference between the loan rate and the deposit rate in bank financing.

In the competitive banking sector, the net interest margin (call it Nim) is given as follows:

$$Nim(\delta_q) = R_L(\delta_q) - R_D(\delta_q) = \frac{M}{p + \delta_b} \text{ where } \delta_b = \frac{\delta_q + \bar{\delta}}{2} \tag{3.16}$$

Proposition 3: *Suppose that the banking sector is competitive. A wider net interest margin may indicate a larger banking sector and greater informational efficiency .*

Proof. The net interest margin increases as the cut-off, δ_q, falls (the banking sector expands). From Proposition 2, we know that $\delta_q < \delta_m$. A fall in δ_q (expansion of bank financing) is clearly a less efficient outcome.

Therefore, a wider net interest margin corresponds to a move away the informationally efficient outcome. □

It is possible that a change in the net interest margin is driven by one of the other parameters in the model. For example, a higher monitoring cost, M, leads to a wider net interest margin and indicates cost inefficiency but does not have implications for informational efficiency.

Traditionally, in empirical work, a change in the bank net interest margin has been interpreted as a consequence of costs in banking. In Proposition 3, I show that it may indicate other forms of efficiency as well. Interestingly, this inefficiency may be corrected fully, as is discussed in Section 4.

3.3.8 Incentive Compatible Monitoring

So far, it has been assumed that monitoring is observable and verifiable, hence contractible. If it is not contractible, a banker monitors only if it is incentive compatible for him to do so. So, the banker monitors only if,

$$(p+\delta_b)(R_L - R_D) - M = (p+\delta_b)\frac{M}{p+\delta_b} - M \geq p\frac{M}{p+\delta_b} \tag{3.17}$$

The banker receives the face value $\frac{M}{p+\delta_b}$. If he monitors and incurs the cost, M, he receives the face value with probability, $(p+\delta_b)$. If he does not monitor, he receives the face value with probability p.

The above condition is always violated and it is never incentive compatible for the banker to monitor the project.

In the present model, there are multiple ways of making monitoring incentive compatible for the banker. One solution (borrowed from Holmstrom and Tirole, 1997) is to require the banker to invest some of his personal wealth in the bank. Suppose that the banker invests $(B - M)$. The incentive problem becomes,

$$(p+\delta_b)\frac{B}{p+\delta_b} - M \geq p\frac{B}{p+\delta_b} \tag{3.18}$$

For $B \geq \frac{M(p+\delta_b)}{\delta_b}$, the banker monitors the project even if he cannot be contractually obligated to do so.

Therefore a banker must invest an amount, $B - M = \frac{pM}{\delta_b}$, of his personal wealth as deposits to credibly signal that he will perform his monitoring duties.

Alteratively, if the banker is wealth constrained, for example, he could be offered a high enough compensation which makes monitoring incentive compatible.

3.4 Policy Implications

Consider that there are two additional types of agents:

There is a benevolent regulator. The regulator aims to maximize social surplus and is unconcerned with distributional effects.

There are also some investors in the equity market. The equity and deposit markets are segmented (Guiso, Haliasos and Jappelli, 2002, Guiso and Sodini, 2013). The participants in the equity market require a higher return than the depositors (due to higher outside options, for example)

3.4.1 Capital Requirements

In this section, I show that if bank equity is more expensive than deposits, it is possible to achieve the efficient outcome using regulatory capital requirements.

A standard assumption in banking literature is that bank equity is more expensive than deposits even after risk-adjustment. Examples include Allen, Carletti and Marquez (2011), Mehran and Thakor (2011). Allen, Carletti and Marquez (2014) present a model in which equity holder earns a higher expected return than depositors as they allow bankruptcy costs to be reduced. Diamond and Rajan (2000) posit that bank capital is socially more expensive than deposits since equity impedes liquidity creation.

Suppose that the cost of equity capital is $k(E) > 0$, with $k(0) = 0$ and $k'(E) > 0$. For simplicity, I assume that following functional form for the cost:

$$k(E) = \beta E, \beta > 0 \tag{3.19}$$

In order to fund the project, the bank will raise $E \in [0, 1]$, in the equity market; the remaining, $(1 - E)$, is raised in the form of deposits. Suppose that the average project that borrows from the bank is of type, δ_b. Including the cost of equity capital, the bank's zero profit condition of the bank becomes:

$$(p + \delta_b)R_L - 1 - M - \beta E = 0 \tag{3.20}$$

If the project succeeds, the borrower repays, R_L, given by:

$$R_L = \frac{1 + M + \beta E}{(p + \delta_b)} \tag{3.21}$$

Lemma 4: *The bank will never voluntarily raise funds using expensive equity.*

Proof. Since lending is competitive, financiers act as Bertrand competitors and contracts are designed to maximize borrower's expected profit, subject to participation constraints.

Therefore, the bank offering the lowest repayment rate (subject to the bank making a zero profit) will capture the entire business. The repayment on loan is minimized by setting $E = 0$. $\square$

Intuitively, since equity is more expensive than deposits, an unregulated bank funds itself entirely with deposits and the banker privately sets $E = 0$.

Suppose that the regulator imposes a capital requirement, $E \in (0, 1]$. Given Lemma 4, the capital requirement always binds for any $E > 0$ and the bank raises E in the equity market to comply with regulation.

Lemma 5: *A regulator may set a capital requirement, $E^R > 0$ such that the equilibrium is efficient ($\delta_q = \delta_m$). E^R is given as follow:*

$$E^R = \frac{(1 + M)(\bar{\delta} - \delta_m)}{2\beta(p + \delta_m)} \tag{3.22}$$

Proof. A project, δ_q, is indifferent between borrowing from the bond market or the bank if,

$$-1 = \delta_q X - (p + \delta_q)\left[\frac{1 + M + \beta E}{(p + \delta_b)}\right] \tag{3.23}$$

By Assumption, $A2$, $\delta_m X = M$. Further, if $\delta_q = \delta_m$, $\delta_b = \frac{\delta_m + \bar{\delta}}{2}$. Substituting in Equation (3.23), the expression for E^R is derived.

If $E < E^R$, then $\delta_q < \delta_m$ and there is scope to improve efficiency by imposing a higher requirement. If $E > E^R$, then $\delta_q > \delta_m$ and there is inefficiency due to under-monitoring. $\square$

Intuitively, costly equity is analogous to a penalty or a tax on bank financing. When a higher penalty is imposed (e.g. through capital requirements), bank financing becomes less attractive and δ_q increases. A benevolent regulator sets $E = E^R$ such that $\delta_q = \delta_m$, which is the efficient outcome.

I show below that the higher efficiency achieved through expensive equity does not necessarily correspond to a smaller net interest margin.

The net interest margin for any E, is given as follows:

$$Nim(E) = R_L(E) - R_D(E) = \frac{(M + \beta E)}{(p + \delta_b(E))} \text{ where } \delta_b = \frac{\delta_q(E) + \bar{\delta}}{2} \tag{3.24}$$

Corollary 1: *Although, the equilibrium is more efficient through the use of expensive equity, the effect on the net interest margin is ambiguous.*

Proof.

$$\frac{dNim}{dE} = \frac{2\beta(2p + \delta_q(E) + \bar{\delta}) - 2\delta_q'(E)(M + \beta E)}{(2p + \delta_q(E) + \bar{\delta})^2} \tag{3.25}$$

$\frac{dNim}{dE} > 0$ only if $\beta > \frac{M\delta_q'(E)}{2p+\delta_q(E)+\bar{\delta}-\delta_q'(E)E}$. □

Equity has two separate effects on the net interest margin. These can be decomposed as follows:

The first effect is that the cost of equity is passed on to the borrower which has a positive effect on the net interest margin (it increases).

The second effect is that higher equity leads to an increase in δ_q (more profitable loans for the bank) which has a negative effect on the net interest margin.

The overall effect depends on which effect dominates. If the cost of equity is high enough (β big enough), the first effect dominates. Otherwise the second effect dominates.

3.4.2 Discussion

In this section, I provide simple policy prescriptions to counter the lending inefficiency which arises in the case of a competitive banking sector. I also point out the limitations.

The way to deal with the inefficiency in bank lending is to affect the cost of borrowing from the bank. As the required repayment, R_L, increases, the equilibrium δ_q increases towards δ_m (higher efficiency). This can be achieved by multiple ways:

1. One way to increase R_L, would be to set a minimum requirement on loan rate. Specifically (in the model), set R_L such that $\delta_q = \delta_m$.

   However, there are practical difficulties in implementing this policy. In the model, I consider only observationally equivalent projects. When projects may be differentiated, the one-rate-fits-all approach will no longer be the case.

   Further, rates imposed conditional on observable qualities of the project is also not a solution to the implementation problem. A lot of the information which a bank uses to evaluate a project is 'soft' (see Rajan, 1992) and may not be observed by external parties, including the regulator.

2. The regulator can impose a capital requirement (or some other tax) on the bank, as described in the previous section. In a competitive lending market, the tax is passed on to the borrower in its entirety. Imposing capital requirements on the bank, gets around the difficulty discussed above and makes the policy implementable.

This policy prescription is reminiscent of De Meza and Webb (1987) in which a higher loan rate is used to achieve the socially efficient level of investment.

3.5 Monopolistic Bank

In this section, I consider a variant of the basic model: the banking sector is monopolistic. The banker extracts some strictly positive profit from lending. The investors in the bond market behave competitively and the market deposits is competitive, as in the previous section.

The problem here is set up differently from the case of the competitive banking sector. In the competitive banking sector, the loan rate is driven down to the point that the bankers make zero profit, in expectation. If a banker sets a higher loan rate, the bank will lose all business to its competitors. This threat does not exist in the monopolistic banking sector. In the monopolistic banking sector, the banker makes the choice of the loan rate to maximize the bank's total expected profits.

An equilibrium consists of a choice of the borrowing contracts by the entrepreneurs and choices of the interest rates by banks and bondholders, which satisfies the following conditions: given equilibrium choices and beliefs of other players, **a.** the banker sets a repayment rate, ${R_L}'$, to maximize his expected profits and **b.** the depositors and marginal borrower (between bond and bank financing) make zero expected profits.

The new loan rate, ${R_L}'$ if not equal to the competitive loan rate, $R_L(\delta_q)$, results in a new cutoff, δ_q' such that projects with $\delta_i \geq \delta_q'$ seek bank financing and projects with $\delta_i < \delta_q'$ seek bond financing. The deposit rate is competitively set at $R_D(\delta_q')$.

For ${R_L}'$ and δ_q' (both parameters are determined in equilibrium), the banker makes an average (per unit loan) expected profit (call it k), given as follows:

$$k = -M + \left(p + \frac{\delta_q' + \bar{\delta}}{2}\right)({R_L}' - R_D(\delta_q')) \tag{3.26}$$

The banker incurs the monitoring cost (at date 0). If the project succeeds (at date 1), the borrower retains the difference between the loan repayment, ${R_L}'$, and the deposit rate, $R_D(\delta_q')$.

The total profit that the banker makes is therefore,

$$\int\limits_{\delta_q'}^{\bar{\delta}} k \,\mathrm{d}\delta_i = (\bar{\delta} - \delta_q')\left[-M + \left(p + \frac{\delta_q' + \bar{\delta}}{2}\right)({R_L}' - R_D(\delta_q'))\right] \tag{3.27}$$

The banker chooses ${R_L}'$ that maximizes the total profit subject to the participation of the marginal borrower, who is of type $\delta_i = \delta_q'$ (if this is satisfied, then participation for all borrowers with $\delta_i > \delta_q'$ is automatically satisfied). The banker's problem is given as follows:

$$\begin{aligned} \underset{R_L{}'}{\text{Max}} \quad & (\bar{\delta} - \delta_q') \left[-M + \left(p + \frac{\delta_q' + \bar{\delta}}{2} \right) (R_L{}' - R_D(\delta_q')) \right] \\ \text{s.t.} \quad & pX - 1 \leq (p + \delta_q')(X - R_L{}') \end{aligned} \tag{3.28}$$

The constraint is that a borrower of type, $\delta_i = \delta_q'$ (weakly) prefers bank financing to bond financing.

Proposition 4: *Suppose that the banker behaves monopolistically. There exists a cutoff, $\delta_i = \delta_q'$, such that an entrepreneur will only seek bank loan if $\delta_i > \delta_q'$ and prefer bond financing if $\delta_i < \delta_q'$. She is indifferent if $\delta_i = \delta_q'$.*

Proof. The argument is identical to the proof for Proposition 1.

To solve for δ_q' and $R_L{}'$, consider the monopolist banker's problem, above. First note that the constraint satisfies with equality. The marginal project of type δ_q' is indifferent between bank and bond financing. To break the indifference, I assume that such a borrower seeks bank financing. Substitute the constraint in the objective function. Take the First Order Condition with respect to $R_L{}'$ to derive the equilibrium values. □

Proposition 5: *Suppose that the banker behaves monopolistically. It is possible that the equilibrium is efficient, i.e., $\delta_q' = \delta_m$. It is also possible that there is under-monitoring, i.e., $\delta_q' > \delta_m$. There is never over-monitoring, $\delta_q' \not< \delta_m$.*

Proof. The marginal project borrowing from the bank is of type δ_q'. From the marginal borrower's participation constraint (equation (3.28), with equality), the required repayment on loan, ${R_L}'$, is derived:

$$ {R_L}' = \frac{1 + \delta_q' X}{p + \delta_q'} \tag{3.29} $$

The deposit rate is set competitively and is given as follows:

$$ R_D(\delta_q') = \frac{1}{\left(p + \frac{\delta_q' + \bar{\delta}}{2}\right)} \tag{3.30} $$

The banker's profit from lending to the marginal borrower, who is of type δ_q', is:

$$ -M + \left[(p + \delta_q'){R_L}' - \left(p + \frac{\delta_q' + \bar{\delta}}{2} \right) R_D(\delta_q') \right] \tag{3.31} $$

The banker incurs the monitoring cost for the project. The marginal borrower succeeds with probability $(p + \delta_q')$ and repays ${R_L}'$. Any depositor is repaid with the average probability of the loan portfolio, $\left(p + \frac{\delta_q' + \bar{\delta}}{2}\right)$.

Substituting, the banker's expected profit from lending to the marginal borrower is:

$$ \delta_q' X - M \tag{3.32} $$

Note that the expected profit from lending to the marginal borrower is 0, for $\delta_q' = \delta_m$. For any $\delta_q' < \delta_m$, the banker makes a negative expected profit. Therefore, the banker sets ${R_L}'$ such that $\delta_q' \geq \delta_m$. $\square$

For a monopolistic banking sector, the net interest margin (call it Nim') is given as follows:

$$Nim'(\delta_q') = R_L{}' - R_D(\delta_q') = \frac{1 + \delta_q' X}{p + \delta_q'} - \frac{2}{(2p + \delta_q' + \bar{\delta})} \tag{3.33}$$

Proposition 5: *Suppose that the banking sector is monopolistic and the equilibrium is inefficient (i.e. $\delta_q > \delta_m$). A wider net interest margin may indicate a smaller banking sector and greater informational inefficiency.*

Proof. The banker chooses $R_L{}'$ to maximize his expected profits.

A lower $R_L{}'$ will lead to a smaller δ_q' (more bank lending). This in turn will result in a higher deposit rate (since $\frac{dR_D}{d\delta_q} < 0$) and ultimately a smaller net interest margin.

It is shown in Proposition 4 that $\delta_q' \geq \delta_m$ when the banking sector is monopolistic. Therefore a smaller δ_q' which is shown to correspond to a smaller net interest margin is a move towards the efficient outcome (which $\delta_q' = \delta_m$).

Using the same arguments, it may be shown that a wider net interest margin corresponds to a bigger δ_q' and a move away the informationally efficient outcome. $\square$

Note that the monitoring cost does not enter the Net interest margin if the banking sector is monopolistic. Intuitively, a higher cost will be absorbed by the monopolistic banker and not passed on to the borrowers. Therefore, the sole source of inefficiency (if there is inefficiency) in this case is monopoly pricing by the banker.

3.6 Conclusion

I have presented a model of external financing choice and lending efficiency. A borrower seeks either monitored bank financing or un-monitored bond financing. I show that the equilibrium with competitive a banking sector is *always* inefficient and is characterized by some degree of over-monitoring. The net interest margin falls with efficiency in lending, supporting the view that the bank interest margin may indicate overall inefficiency in the credit allocation environment (as opposed to simply cost inefficiency in bank lending). In terms of policy implications, imposing a higher bank loan rate (for example, via costly capital requirements) will increase the efficiency in credit allocation across the banking sector and the bond market.

Finally, I show that in contrast to the competitive equilibrium, if the banking sector is monopolistic, the equilibrium may be either efficient or it may be inefficient and characterized by some under-monitoring.

I have considered two separate cases: the competitive and the monopolistic banking sector. I have been silent on the welfare implications of a transition from one regime to another. This is a potentially interesting direction for future research.

3.7 Appendix

Derivation of δ_q (competitive banking sector):

The indifference condition between bank and bond financing is written as,

$$pX - 1 = (p + \delta_q)[X - R_L(\delta_q)]$$

Substituting $R_L(\delta_q)$,

$$pX - 1 = (p + \delta_q)\left[X - \frac{2(1 + M)}{2p + \delta_q + \bar{\delta}}\right]$$

The equation is quadratic in δ_q and is rearranged as,

$$\delta_q^2 X + \delta_q[(2p + \bar{\delta})X - (2M + 1)] - (2pM - \bar{\delta}) = 0 \tag{3.34}$$

A real solution exists only if the determinant ≥ 0,

$$[(2p + \bar{\delta})X - (2M + 1)]^2 + 4X(2pM - \bar{\delta}) \geq 0$$

If a real solution exists,

$$\delta_q = \frac{1}{2X}\left[-((2p + \bar{\delta})X - (2M + 1)) \pm \sqrt{((2p + \bar{\delta})X - (2M + 1))^2 + 4X(2pM - \bar{\delta})}\right]$$

There are three separate cases to consider:

1. Both roots of δ_q are positive. In this case either root may be the equilibrium.

2. One root is positive and the other root is negative. Then δ_q equals the positive root.

3. Finally if both roots are negative, bank financing dominates bond financing for all projects, irrespective of type. So, we set δ_q equals 0.

Bibliography

ALLEN, F. (1990): "The Market for Information and the Origin of Financial Intermediation," *Journal ofFinancial Intermediation*, 1, 3–30.

ALLEN, F., AND E. CARLETTI (2013): "Deposits and Bank Capital Sructure," *Working Paper*.

ALTMAN, E., A. GANDE, AND A. SAUNDERS (2010): "Bank Debt versus Bond Debt: Evidence from Secondary Market Prices," *Journal of Money, Credit and Banking*, 42(4), 755–767.

BECK, T., A. DEMIRGUC-KUNT, AND R. LEVINE (1999): "A new Database on Financial Development and Structure," World Bank.

BERNANKE, B. (1983): "Nonmonetary Effects of the Financial Crisis in the propagation of the Great Depression," *American Economic Review*, 73(3), 257–276.

BESANKO, D., AND G. KANATAS (1993): "Credit Market Equilibrium with Bank Monitoring and Moral Hazard," *Review of Financial Studies*, 6(1), 213–232.

BHATTACHARYA, S., AND G. CHIESA (1995): "Proprietary Information, Financial Intermediation and Research Incentives," *Journal of Financial Intermediation*, 4(4), 328–357.

BOLTON, P., AND X. FREIXAS (2000): "Equity, Bonds and Bank Debt: Capital Structre and Financial Market Equilibrium under Asymmetric Information," *Journal of Political Economy*, 108(2), 324–351.

Boot, A. (2000): "Relationship Banking: What do we Know?," *Journal of Financial Intermediation*, 9, 7–25.

Boot, A., and A. Thakor (1997a): "Financial System Architecture," *Review of Financial Studies*, 10(3), 693–733.

——— (1997b): "Banking Scope and Financial Innovation," *Review of Financial Studies*, 10, 1099–1131.

Boyd, J., and E. Prescott (1986): "Financial Intermediary Coalitions," *Journal of Economic Theory*, 38(2), 211–232.

Chemmanur, T., and P. Fulghieri (1994): "Reputation, Renegotiation and the Choice between Bank Loans and Publicly Traded Debt," *Review of Financial Studies*, 7(3), 475–506.

DeMeza, D., and D. Webb (1987): "Too Much Investment: A Problem of Asymmetric Information," *Quarterly Journal of Economics*, 102(2), 281–92.

Demirguc-Kunt, A., and H. Huizinga (1999): "Determinants of Commercial Bank Interest Margins and Profitability: Some International Evidence," *World Bank Economic Review*, 13(2), 379–408.

——— (2000): "Financial Structure and Bank Profitability," World Bank Policy Research Working Paper.

Diamond, D. (1984): "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51, 393–414.

——— (1991): "Monitoring and Reputation: The Choice between bank loans and directly placed debt," *Journal of Political Economy*, 99(4), 689–721.

Diamond, D., and R. Rajan (2000): "A Theory of Bank Capital," *Journal of Finance*, 55, 2431–2465.

Fama, E. (1985): "What's Different About Banks?," *Journal of Monetary Economics*, 15(1), 29–39.

Gande, A., and A. Saunders (2012): "Are Banks Still Special When There is a Secondary Market for Loans?," *Journal of Finance*, 67(5), 1649–1684.

Guiso, L., M. Haliassos, and T. Jappelli (2002): in *Household Portfolios*Cambridge, MA. MIT Press.

Guiso, L., and P. Sodini (2013): "Household Finance: an Emerging Field," in *Handbook of Economics and Finance*, ed. by G. Constantinides, M. Harrris, and R. Stulz, pp. 1397–1532, Amsterdam, North Holland.

Holmstrom, B., and J. Tirole (1997): "Financial Intermediation, Loanable funds and the real Sector," *Quarterly Journal of Economics*, 112(3).

James, C. (1987): "Some evidence on the Uniqueness of Bank Loans," *Journal of Financial Economics*, 19, 217–238.

Leland, H. E., and D. H. Pyle (1977): "Information Asymmetries, Financial Structure and Financial Intermediation," *Journal of Finance*, 32, 371–387.

Morellec, E., P. Valta, and A. Zhdanov (2013): "Financing Investment: The Choice between Bonds and Bank Loans," *Working Paper*.

Rajan, R. (1992): "Insiders and Outsiders: The Choice between bank loans and directly placed debt," *Journal of Finance*, 47, 1367–1400.

Ramakrishnan, R., and A. Thakor (1984): "Information Reliability and a theory of Financial Intermediation," *Review of Economic Studies*, 51(3), 415–432.

Repullo, R., and J. Suarez (2000): "Entrepreneurial Moral Hazard and Bank monitoring: a model of the Credit channel," *European Economic Review*, 44(10), 1931–1950.

Salop, S. (1979): "Monopolistic Competition with Outside Goods," *Bell Journal of Economics*, 10(1), 141–156.

Sapienza, P. (2002): "The effects of Banking Mergers on Loan Contracts," *Journal of Finance*, 57(1), 329–367.

Song, F., and A. Thakor (2010): "Financial System Architecture and the Coevolution of Banks and Capital Markets," *Economic Journal*, 120(547).

Townsend, R. (1979): "Optimal Contracts and Competitive Markets with Costly State Verification," *Journl of Economic Theory*, 21, 265–293.

von Thadden, E.-L. (1995): "Long-Term Contracts, Short-term Investment and Monitoring," *Review of Economic Studies*, 62(4), 557–575.

Yosha, O. (1995): "Information Disclosure Costs and the choice of Financing Source," *Journal of Financial Intermediation*, 4(1), 3–20.

Chapter 4

Is Bank Equity more expensive than Deposits?

4.1 Introduction

A standard assumption in banking literature is that bank equity is costlier than deposits. The assumption drives many results but it lacks sound theoretical foundations (Admati, et. al., 2011). Furthermore, there is strong empirical evidence that cognitive abilities affect investor decision regarding participation in equity markets. Overwhelming majority of investors hold bank deposits but only higher IQ investors own equity (references below). We propose a simple model which provides a joint explanation for why the return on bank equity is higher than that on deposits (although there is universal risk neutrality) and why deposits and equity are held by different groups of individuals.

Standard models in Finance (e.g. CAPM) predict that investors should invest in the market portfolio (which spans all asset classes, by definition). However, there is considerable evidence that the market for deposits is significantly segmented from equity markets (Guiso, Haliassos and Jappelli, 2001 and Guiso and Sodini, 2013). Significant fractions of investors only hold information-insensitive assets such as bank savings accounts and they refrain from participating in equity markets. While participation costs provide some explanation for this phenomenon, the size of direct participation costs (such as fees) do not satisfactorily explain the degree of non-participation in equity markets. It is also unclear why participation costs will deter the affluent from investing the equity markets.

We consider a model in which potential investors have different monitoring skills. There are two types of investors: skilled and unskilled, and the type is public knowledge. The skilled can monitor a project and increase its productivity, whereas the unskilled cannot. Banks emerge endogenously and the bank manager is always a skilled investor. In a coalition, the skilled banker may divert a fraction of the returns from the unskilled and this diversion cannot be detected by the unskilled.

If diversion is large enough, bank capital structure is relevant and the optimal arrangement entails that the unskilled investors become depositors and the skilled, equity-holders. An unskilled investor does not participate in an all-equity bank because if they own a security of the same seniority as the skilled, the payoff (after diversion) does not meet their outside option (so, they prefer to invest on their own than have equity in a bank). To attract funds from the unskilled, the banker offers them a senior deposit contract and retains the residual claim for himself and the other skilled investors.

Therefore, we have here a segmented deposit and equity markets. This arises endogenously in our setting as the seniority of deposits (relative to the equity held by the skilled investors) ensures that the unskilled investor's participation constraint is satisfied.

As a result of the market segmentation, bank equity is more expensive than deposits. The skilled investors (equity) have a higher outside option than the unskilled (deposits). Further, since skill is scarce, the skilled investors retain the surplus when they manage the funds of the unskilled. Thus, in this setting the higher return on equity is not related to risk. Instead, it is a premium for skill which is scarce.

The skilled equity-holders face the following tradeoff: on the one hand, there is economies of scale in a coalition. On the other hand, equity-holders incur a cost to prevent diversion by the banker. The trade-off determines the size of bank's equity capital. Further, the degree of diversion along with the productivity differential between skilled and unskilled (diversion constraint) determines the optimal leverage ratio of the bank. Taken together, we have implications for the optimal bank size.

We provide some interesting and potentially testable empirical implications: First, our model predicts that banks become larger and more levered as the legal environment becomes stronger. Secondly, banks become larger and more levered as the deposit market becomes less competitive. Finally, banks become larger and better capitalized (holding leverage constant) as the monitoring technology improves.

4.2 Literature

The present paper is related most closely to two separate strands of the literature: i. Cognitive ability and stock market participation and ii. Corporate Finance of Banks.

4.2.1 Cognitive ability and stock market participation:

There is an emerging strand of literature which suggests that participation in equity markets correlates with investor cognitive abilities. More intelligent investors are more likely to own stocks.

Grinblatt, Keloharju and Linnainmaa (2011) find that equity market participation increases in IQ, controlling for wealth, income, age and other demographic and occupational information. They find that lack of cognitive skills is so fundamental as a driver of non-participation that it defers large amounts of wealth from entering the equity market. Further, they find that IQ plays an equally important role with regards to participation in a sub-sample when only the affluent are considered.

Cole, Paulson and Shashtry (2014) find that an additional year of education (which likely affects cognitive abilities positively) leads to a 4 percentage points increase in probability that an individual owns equities. They also find more direct evidence of cognitive ability positively affecting equity market participation by studying siblings, who grew up with similar backgrounds. Van Rooji, Lusardi and Alessie (2011) find that equity ownership increases sharply with financial literacy and education, in general.

Christelis, Jappelli and Padula (2010) find that propensity to invest in equities is strongly associated with cognitive abilities (measured by mathematical, verbal and recall tests). Further, propensity to hold informationally insensitive assets such as bank savings accounts is not affected by cognitive abilities.

4.2.2 Corporate Finance of Banks:

A recent paper that is closely related to ours is Allen and Carletti (2013). As we have here, the expected return on bank equity is greater that the expected return on deposits, in a risk neutral setting. However, there are important differences:

1. Allen and Carletti (2013) assume that the deposit and equity markets are segmented; without market segmentation depositors could invest directly in equity and the return differential will no longer exist. We derive the market segmentation as an endogenous result.

2. An essential ingredient in Allen and Carletti (2013) is costly bankruptcy. Equity is more expensive than deposits as it adds value through reducing inefficient bankruptcy costs. We do not use costly bankruptcy for our results.

Hart and Moore (1995) analyze the role of long term debt in restricting the manager from raising new funds to invest in negative NPV projects. Similar to our story, the debt contract (which is a long-term contract) in Hart and Moore (1995) curbs how much the manager may divert for private benefits (for example, empire building). Although we are silent on maturity structure of debt (since we have a one-period model), we have additional implications with regards to returns on different securities.

Diamond (1984) uses a diversification argument to justify the use of the debt/deposit contract in the bank capital structure. As the number of depositors go to infinity, stochastic returns across uncorrelated projects become deterministic and the banker can offer risk-less payoffs to depositors. The debt contract is optimal in this model and there is no equity capital. Similar to the argument in Diamond (1984), the coalition in our model benefits from economies of scale. In contrast to Diamond (1984) in which the optimal contract is debt, here the bank optimally holds both deposits and outside equity; capital structure is relevant.

Calomiris and Kahn (1991) derive conditions under which demandable debt (with sequential service constraint) is optimal in banking. To restrict the banker from absconding with the banks' profits in the bad state, uninsured depositors monitor the banker and pull out their deposits (run) when they fear absconding. The sequential service constraint circumvents the free rider problem of small depositors in incurring monitoring costs. Monitoring depositors are first in line to withdraw if they get a negative signal. A free riding depositor is later in the queue and hence has a lower expected payoff upon withdrawal. A sequential service constraint thus preserves the monitoring incentives of depositors. Like Diamond (1984), there is no equity capital in this model.

Diamond and Rajan (2000, 2001) provide a theory of bank capital structure where they differentiate between the demandable debt contract (sequential service constraint) and capital (standard debt or equity) by their renegotiability. The banker may threaten to withhold his specific skills and extract some rent from capital. Demandable debt gets around the hold-up problem because if the banker attempts to renegotiate the terms of the demandable debt contract, depositors run on the bank and pull out their deposits; this disintermediates the banker and destroys his rents. Thus, the renegotiable capital sacrifices liquidity creation (loan-making) to provide stability in the poor state. Equity capital is expensive in this model from a social perspective as the equity claim prevents liquidity creation. In our model, we look at the cost of bank equity vis-a-vis bank deposits in the standard sense of required returns on these securities.

Holmstrom and Tirole (1997) study a model with effort moral hazard in which internal capital strengthens monitoring incentives and allows external funds to be raised. They do not consider security design issues and in their model debt and equity are identical.

In Morrison and White (2005) the sound investors (similar to our skilled investors) form the bank's equity capital and the unsound investors are depositors. We use this idea of different types of agents (set apart by skill and hence, outside options). However, their focus too is not optimal security design and in fact, they assume the debt contract and do not derive it endogenously. Indeed, it is easy to show that the contract they refer to as the deposit contract in their model could equally be an equity contract.

Allen, Carletti and Marquez (2011) show that improved monitoring incentives (due to higher capital) attracts more deposits and hence surplus extraction. Mehran and Thakor (2009) essentially reinforce the same point (capital improves monitoring incentives). They have a direct channel that higher bank capital leads to a higher survival probability of the bank (at an interim date) which increases the marginal benefit of monitoring. Further, this effect is magnified as the higher monitoring increases the value of the relationship loan portfolio.

In our model, higher equity capital attracts more deposits by easing the diversion constraint. The argument does not rely on improved monitoring incentives and therefore highlights a separate channel of why higher capital may lead to more deposits.

Our theory provides a link between bank capital structure and the strength of the legal system. Diamond (2004) presents a model of optimal debt when contract enforcement is ineffective and expensive. If the lender goes to the court, it imposes a cost on the overall value of the firm. Thus, ex-ante a lender cannot commit to go to the court even if he observes borrower misbehavior. A senior claim (such as debt) solves this problem of lender passivity as the entire cost of going to the court is borne by the borrower.

We also have implications for the optimal size of bank. Cerasi and Daltung (2000) derive the optimal size of a bank by trading off the costs and benefits of a larger (diversified) loan portfolio. Diversification increases the banker's incentive to monitor the loans. However, monitoring more loans also entails overload costs (since there is a limit to the number of loans the banker can monitor). In our model, the trade-off for the optimal size comes from the liabilities side of the balance sheet as opposed to the loan portfolio.

4.3 Model

4.3.1 Set-up

We consider a one-period economy in which all agents are risk-neutral. At $t = 0$ there is investment in a project and at $t = 1$ returns are realized. All agents consume at time,$t = 1$. Each agent has an initial endowment of 1 unit at $t = 0$. There are many projects to be undertaken, and if undertaken, a project yields either X (success) or 0 (failure) per unit of investment at $t = 1$. If the project is monitored, it returns X with probability p_h and 0 with probability $(1 - p_h)$. If not monitored, the success probability is p_l, where $p_l < p_h$. Monitoring the project requires exerting effort which implies a non-pecuniary cost, F, where $F \geq 0$. We assume that monitoring is efficient in the sense that the increase in the expected return exceeds the monitoring cost:

A1: $(p_h - p_l)X > F$

We also assume that all projects have positive net present value (NPV) even in the absence of monitoring:

A2: $p_l X - 1 > 0$

There are two types of agents: skilled and unskilled. The agent type is public information. Only the skilled are able to monitor projects, while the unskilled investors lack the expertise to monitor. There are N agents in total with a proportion λ of them being skilled and the rest unskilled.

The outside option of each agent is determined by the expected return of investing on his own. Because monitoring is efficient (Assumption 1), the outside option of the skilled agents ($p_h X - F - 1$) is greater than that of the unskilled ($p_l X - 1$).

The key friction in our model is diversion. If a skilled agent manages the funds (the manager) of unskilled, he may divert a fraction ϕ of the realized output. The diverted amount can not be verified in a court of law. Further, the manager cannot credibly commit to not divert funds from the unskilled even if it is beneficial for him to do so ex-ante. A skilled manager can divert from unskilled but not from other skilled agents. Finally, an unskilled manager cannot divert.

The skilled agents are scarce relative to the unskilled:

A3: $\lambda \leq \frac{p_l - \phi p_h}{p_l}$

Hence, the skilled investors keep any surplus that they generate from managing the funds of the unskilled. This assumption may be relaxed to the extent that the skilled investors get a non-zero fraction of the surplus (incremental profitability due to monitoring).

Because of the increase in expected return of a project due to monitoring exceeds the monitoring cost, maximization of the net social surplus (efficiency) requires that all funds invested are managed by skilled agents.

In this setting agents decide whether they will invest on their own or coalesce with other agents to invest jointly. One of the coalition members becomes the manager. The coalition of agents is the bank and the manager the banker. The banker may be a skilled or an unskilled agent. In order to make our results clearer, we assume that the monitoring cost, F, equals 0.

4.3.2 Some General Results

Consider an unskilled agent who manages a bank with n members. The expected return of the combined investment of n units is np_lX. The unskilled manager can credibly promise to each bank member a maximum of p_lX. A higher offer is not credible as it violates the participation constraint of the manager or some of the other bank members. A skilled agent has a higher outside option ($p_hX - F > p_lX$ from investing on his own) and so will never delegate investment to an unskilled agent. An unskilled agent may be indifferent between investing on his own and delegating to another unskilled agent. Hence,

Lemma 1: *If the banker is an unskilled agent, no skilled agent will join the bank.*

Corollary: *Banks managed by an unskilled agent may only be formed by unskilled agents.*

Furthermore, because an unskilled manager cannot divert part of the realized return, the other unskilled bank members will receive the promised repayment, X, if the project succeeds regardless of the security issued by the banker. That is,

Lemma 2: *If the banker is an unskilled agent, the bank capital structure is irrelevant.*

If, instead, the banker is a skilled agent, the expected return of the combined investment of n units is np_hX. The skilled manager can credibly promise a maximum of p_hX to each bank member. Thus, a skilled agent is indifferent between investing on his own and delegating to another skilled. A skilled manager can attract an unskilled agent by offering him $p_lX + \epsilon$, where ϵ can be infinitesimally small. Hence,

Lemma 3: *If the banker is a skilled agent, both skilled and unskilled agents may join the bank.*

4.3.3 Benchmark Case

In the absence of diversion, all unskilled agents would like to join a bank managed by a skilled agent. However, because a skilled banker can divert part of the realized return there is an upper bound to the number of unskilled agents who can receive the promised repayment if they join a bank managed by a skilled banker. As we show below, the extent of diversion will determine the type of security the banker issues and the number of unskilled agents who join the bank per skilled bank member.

We begin by considering the case where the only skilled agent in the bank is the banker and the security issued is equity. At $t = 0$, the skilled banker accepts outside funds from I_u unskilled investors. The total funds available to the bank, including the banker's endowment, are (I_u+1). The expected return to the bank from investing this amount is $p_h(I_u+1)X$. The skilled banker diverts a fraction $(1-\phi)$ of the realized cash flow (non-verifiable). The fraction that remains is distributed among all the bank members (including the banker), as they have a claim of the same seniority (equity claim). Given that the banker cannot credibly commit to not divert, the expected payoff to each unskilled member is $\phi p_h X$.

If the expected payoff to the unskilled meets their outside option, $p_l X$, diversion by the skilled is harmless and it does not matter in terms of expected payoffs whether the unskilled members in the coalition hold equity (as above) or some other claim, like debt [1]. So, the bank's capital structure is not relevant in this context. We do not consider this case in the rest of the paper. Formally, from now on we assume that:

A3: Diversion is large enough such that $\phi < \frac{p_l}{p_h}$.

If this is the case, the equity claim no longer meets the outside option of the unskilled agent as $\phi p_h X < p_l X$. The unskilled investors do not participate in the all-equity bank. From the perspective of the skilled investors this is not desirable as attracting funds from the unskilled allows them to extract surplus. Further, from an efficiency point of view, it is desirable that the maximum possible amount of funds is invested through the skilled agents and therefore, this is an inefficient outcome in terms of the net social surplus. The above discussion leads us to the following Lemma:

Lemma 4: *Suppose that the banker diverts a fraction,* $(1 - \phi)$ *of the cash flows, where* $\phi < \frac{p_l}{p_h}$*. Unskilled investor will not participate in the bank as equity holders and prefer to invest individually.*

[1] The security held by the banker cannot be senior to that issued to the unskilled bank members.

4.3.4 Optimal Contracts

In this section we derive the optimal contract for each type of agents. The banker maximizes his profit subject to two constraints: i) the participation constraint of the unskilled investors, and ii) the constraint that the total verifiable cash flow (after diversion) must be weakly greater than the total promised payments to the unskilled bank members and the banker (the diversion constraint). The contract offered to the unskilled bank members specifies the repayment, R_u, if the project succeeds (and possibly its seniority relative to the banker's claim). The banker's payoff is R_s. If the project fails, all bank members receive zero. Formally, the banker solves the following problem:

$$\begin{aligned} \underset{I_u}{\text{Max}} \quad & p_h(I_u+1)X - p_h R_u I_u - F \\ \text{subject to} \quad & \phi p_h(I_u+1)X \geq p_h R_u I_u + p_h R_s \\ & p_h R_u = p_l X \end{aligned}$$

First note that the banker's profit function is strictly increasing in the number of unskilled agents joining the bank, I_u. This is true because, by Assumption 2, the banker retains the full incremental profitability (due to monitoring), $(p_h X - p_l X)$, of each unskilled investor's funds and the promised repayment equals the unskilled outside option. That is, the unskilled participation constraint is binding. By substituting the unskilled participation constraint into the diversion constraint, we get:

$$(p_l X - \phi p_h X) I_u \leq \phi p_h X - p_h R_s \tag{4.1}$$

Given Assumption 4, $\phi < \frac{p_l}{p_h}$, the LHS increases with the number of unskilled agents joining the bank, Iu. Hence, in order to maximize I_u consistent with the diversion constraint being satisfied we set $R_s = 0$. Also, because the banker's profit increases in I_u, the diversion constraint will always be binding. This allows us to determine the number of unskilled agents joining the bank, which is $I_u = \frac{\phi p_h}{p_l - \phi p_h} = d$. Henceforth, we refer to the amount of funds provided by the unskilled in the bank as d.

The optimal arrangement entails that the unskilled investors have priority over the verifiable fraction of the cash-flow and so they receive the most senior claim which can be interpreted as debt (deposit). In fact, it is risky debt with face value $R_u = \frac{p_l X}{p_h}$. This credibly ensures that the unskilled investors in the bank earn their outside option (as opposed to the all-equity bank). The banker becomes the bank equity holder (residual claimant).

If a skilled investor joins another, the benefit is cost sharing (as in Diamond, 1984) since the monitoring cost is fixed for any scale of the project. However, in our model, this effect is interesting from a security design perspective. In Diamond (1984), the optimal contract is the debt contract. Here, we show that the optimal contract for the skilled investor is the equity contract.

The following Lemma summarizes these results:

Lemma 4: *The optimal contract for the unskilled is debt (deposit) and the skilled banker holds the residual claim (equity).*

4.3.5 More than One Skilled Bank Member

Now, we consider the case in which other skilled agents may join the bank. Let us denote the total number of skilled agents joining the bank (including the banker) by k and the total number of unskilled agents by d. From the above discussion it is clear that the real constraint for the banker in his attempt to maximize his profits is the diversion constraint and there is a maximum amount of debt consistent with this constraint being satisfied. Also, a skilled agent can monitor the banker regarding diversion. Hence, the optimal contract for any skilled bank member cannot be as senior as the one given to the unskilled agents because that would make the diversion constraint more binding. Thus, the optimal contract to the skilled investor should be junior to debt. This contract could be equity, preferred equity or even debt junior to the debt given to the unskilled. For simplicity, from now on we refer to this contract as equity. In this case, the banker solves the following problem:

$$\begin{aligned} \underset{d,k}{\text{Max}} \quad & \frac{1}{k}[p_h(d+k)X - p_h R_u d - F] \\ \text{subject to} \quad & \phi p_h(d+k)X \geq d p_h R_u + k p_h R_s \\ & p_h R_u = p_l X \end{aligned}$$

As in the case of one skilled bank member, the banker's profit function is strictly increasing in the number of unskilled agents joining the bank, d. Also, by substituting the unskilled participation constraint into the diversion constraint, we obtain:

$$(p_l X - \phi p_h X)d \geq \phi p_h X k - k p_h R_s \tag{4.2}$$

Because, by Assumption 4, $p_l - \phi p_h > 0$ and the banker profit increases in d, profit maximization requires $R_s = 0$ and the diversion constraint being binding. That is, at the optimum:

$$(p_l X - \phi p_h X)d = \phi p_h X k$$
$$\Rightarrow \left(\frac{d}{k}\right)^* = \frac{\phi p_h}{p_l - \phi p_h} \tag{4.3}$$

Equation (4.3) gives us the optimal leverage.

The intuition is straightforward. Given the level of equity capital in the bank and the level of diversion, $1 - \phi$, the banker can credibly promise to depositors only up to the amount of output which can not be diverted, $\phi p_h X$. This, in turn, determines the maximum amount of funds that the depositors are willing to provide to the bank.

We summarize the above discussion in the following proposition:

Proposition 1: *Both skilled and unskilled investors participate in a bank:*

1. There are d unskilled depositors and,

2. The bank's equity capital is made of k skilled investors, one of whom manages the bank.

Proposition 2: *Bank equity is more expensive than deposits.*

Proof. The overall return to the investments made by the bank is $(d + k)p_h X - F$. The banker incurs the monitoring cost and each unit invested yields a return, $p_h X$.

The depositors are the unskilled investors and they earn their outside option, $p_l X$.

The equity claim is held by the skilled investors. Since skill is scarce, the skilled investors keep the entire surplus. Including the monitoring cost, the net payoff to the skilled investor is $p_hX + \frac{d}{k}(p_h - p_l)X - \frac{F}{k} > p_lX$. $\square$

As long as skill is scarce, the expected return on equity is strictly greater than the expected return on deposits. Thus, in this risk-neutral setting the higher return on equity is not related to risk. Instead, it is a premium for skill which is scarce. Therefore, both the assumptions of diversion and skill scarcity are required for the result that equity is more expensive than deposits.

The banker's profit function is increasing in k, as a higher level of equity is beneficial from a cost sharing perspective. Therefore, all skilled agents coalesce to form a single bank and,

$$k^* = \lambda N \tag{4.4}$$

Lemma 5: *There is a unique Nash equilibrium in which there is a single bank in which size of equity is unbounded with, $k^* = \lambda N$, where λN is the number of skilled investors in the world.*

4.3.6 Extension

In the extension we exogenously introduce a cost for skilled investors in the coalition; the peer-monitoring cost. With the inclusion of the peer-monitoring cost, most of the above results still hold, except that the size of bank equity is bounded.

The banker may divert from a skilled bank member, unless the skilled bank member monitors the banker at a positive cost. We assume that this cost is increasing in the size of the coalition. As the bank's balance sheet expands, the skilled bank member needs to incur a higher peer-monitoring cost to prevent diversion.

Further, note that the leverage ratio is fixed, as the diversion constraint binds. Using the optimal leverage ratio, the size of the bank is,

$$d + k = \frac{\phi p_h}{p_l - \phi p_h} k + k \tag{4.5}$$

The size is captured in terms of the exogenously given parameters, (ϕ, p_h and p_l) and k, which is determined within the model. Therefore the peer monitoring cost which is a function of the bank size can be written as a function of bank equity only. For simplicity, we consider a linear cost function and denote the peer monitoring cost as γk per skilled member in the bank.

The banker maximization problem is rewritten to include the peer-monitoring cost:

$$\begin{aligned}
&\underset{d,k}{\text{Max}} && \frac{1}{k}[p_h(d+k)X - p_h d R_u - F] - k\gamma \\
&\text{subject to} && \frac{d}{k} = \left\{\frac{d}{k}\right\}^* = \frac{\phi p_h}{p_l - \phi p_h} \\
& && p_h R_u = p_l X
\end{aligned}$$

While a higher level of equity is still beneficial from a cost sharing perspective, it also imposes a higher peer-monitoring cost; the optimal level of equity is determined at the margin. Substituting the constraints into the profit function and taking the first order condition with respect to k, we derive the optimal size of bank equity, k^*:

$$k^* = \sqrt{\frac{F}{\gamma}} \tag{4.6}$$

Lemma 6: *There is a unique Nash equilibrium in which there are $\frac{\lambda N}{k^*}$ banks and each bank has equity, $k^* = \sqrt{\frac{F}{\gamma}}$, where λN is the number of skilled investors in the world.*

Proof. First consider the case that $k^* = 2$. We show that there is a unique Nash equilibrium in which there are $\frac{\lambda N}{2}$ banks each with $k = 2$. Suppose that we start with λN banks with $k = 1$. Then any banker will profitably deviate by coalescing with any other and form a bank which is twice as big the original banks and this is a strictly profitable strategy for the skilled investors (the unskilled investors are unaffected). Similarly, if we start with 1 bank with N skilled members, they profitably deviate to break up into smaller banks till eventually there are $\frac{\lambda N}{2}$ banks each with $k = 2$.

Now, we consider the case that $k^* \geq 3$. There is a unique Nash equilibrium in which there are $\frac{\lambda N}{k}$ banks each with $k = k^*$. We need to rule out the possibility that there is a coordination failure problem. Suppose that we start with λN banks with $k = 1$; each skilled investor manages his own bank. Consider then, a banker deviates by coalescing with another and forming a bank twice as big as the original banks. Is this a profitable deviation? The benefit to coalescing is $B = F - \frac{F}{k}$ and the cost is $C = k\gamma$. Coalition of skilled investors is possible if and only if benefits to doing so exceeds the cost, $B > C$. (Due to the concavity of B and the linearity of C in k) If $B > C$ for a certain size k^*, it must be true that $B > C$ for any $k < k^*$. Thus, there is no coordination failure problem and bank equity, k^* is achieved as a unique Nash equilibrium. □

Lemma 7: *There is a unique optimal size of the bank which is bounded. It is given* *as follow:*

$$(d+k)^* = \frac{\phi p_h \sqrt{\frac{F}{\gamma}}}{p_l - \phi p_h} + \sqrt{\frac{F}{\gamma}} \tag{4.7}$$

Proof. The expression for the optimal size of the bank is readily obtained by combining the results in Proposition 2 and Lemma 6. □

Intuitively speaking, the sound investors who are the equity-holders in the bank trade off the costs and benefits of being in the bank. The optimal k^* determines the participation of unskilled depositors, d^* in the bank via the optimal leverage ratio given in Proposition 2 (which comes from the diversion constraint).

4.4 Robustness

4.4.1 Bargaining Power

Till now we have considered the case that skilled have all the bargaining power and keep the full surplus from coalescing with the unskilled (since skill is scarce, $A3$). To relax the assumption, denote the bargaining power of the skilled by $\gamma \in [0, 1]$. The skilled keep a fraction γ of the surplus generated using the funds of the unskilled.

Suppose that there are d skilled and k unskilled investors in the bank. Consider the case $F = 0$.

The surplus generated from investing 1 unit of unskilled funds by the skilled is $(p_h - p_l)X$. A depositor extracts a fraction $(1 - \gamma)$ of the surplus. A depositor receives a payoff:

$$p_l X + [(1-\gamma)(p_h - p_l)X] \tag{4.8}$$

Similarly, it may be worked out that the skilled investor has a payoff,

$$p_h X + [\gamma \frac{d}{k}(p_h - p_l)X] \tag{4.9}$$

For any $\gamma > 0$, the return on equity is higher than the return on deposits. For $\gamma = 0$, both equity and deposits have the same payoff, $p_h X$.

Under no circumstances, equity has a lower payoff than deposits. If that were the case, the skilled investors will choose to become depositors (instead of being equity holders) until the payoffs to deposit and equity are at least equal.

The diversion constraint becomes:

$$\begin{aligned} \phi p_h(d+k)X &= d[p_l X + (1-\gamma)(p_h - p_l)X] \\ \Rightarrow \frac{d}{k} &= \frac{\phi p_h}{(1-\phi)p_h - \gamma(p_h - p_l)} \end{aligned} \tag{4.10}$$

4.4.2 Deposit Insurance

Suppose that the regulator operates a deposit insurance scheme such that $d_r(k)$ depositors are fully insured in the high state.

To fund the scheme the regulator taxes the bank equity holders. The maximum tax that may be levied by the regulator is the verifiable portion of the bank's cash flow after diversion, which is given by $\phi(d_r + k)p_h X$. The tax is distributed among the d_r depositors and each depositor receives $\phi(d_r + k)p_h X/d_r$. This amount should meet the outside option of the unskilled investors to ensure their participation in the bank. That is,

$$\begin{aligned} \frac{1}{d_r}\phi(d_r + k)p_h X &= p_l X \\ \Rightarrow \frac{d_r}{k} &= \frac{\phi p_h}{p_l - \phi p_h} \end{aligned} \tag{4.11}$$

Therefore if the regulator operates a balanced budget deposit insurance scheme, the outcome is identical to the case with no deposit insurance.

The same arguments hold for partially insured deposits.

The mechanism analyzed above only insures the depositors in the high state. Full insurance, regardless of state will require the regulator to tax the bank at the time of formation (date 0) and store the amount till the next date. This is socially sub-optimal as investment in a project (whether monitored or not) is more efficient than simply storing the funds (by Assumption).

To summarize,

Proposition 4: *There exists no balanced-budget (accurately priced) deposit insurance scheme which will ease the diversion constraint relative to the unregulated case.*

4.5 Implications

4.5.1 Policy

Myers and Majluf (1984) show that equity issuance will be more expensive than debt for current owners of the firm in a market characterized by asymmetric information (the lemon's problem). Admati et al (2011) conjecture that this cost only applies if equity issuance is by a discretionary decision of the bank. If, instead, a higher capital requirement is imposed by the regulator, new equity issued will not be subject to the adverse selection cost.

However, in our model if capital requirement binds (i.e. if capital requirement is higher than the privately optimal level), the outcome is socially costly. Given that skilled investor manages the bank and unskilled investor deposits, higher leverage means more productive use of funds. If a policy restricts leverage (higher capital requirement), this is welfare reducing.

4.5.2 Predictions

We can derive some interesting and potentially testable predictions:

1. *Bank becomes larger and more levered as diversion decreases (higher ϕ) (From Equations (4.3) and (4.7)).*

   Intuitively, as diversion decreases, lower levels of equity needs to be in place to convince the unskilled investors to accept the deposit contract. Further, given that size of equity is determined by equation ., the lower leverage leads to a smaller bank.

One interpretation is that the parameter ϕ captures the quality of the legal system. The prediction is that as the legal system becomes stronger (contract enforcement improves), banks become larger and more levered.

2. *Bank becomes larger and more levered as skilled have higher bargaining power (higher γ). (From Equation (4.10))*

   As the bargaining power of the skilled investor increases (the market for deposit becomes less competitive) the skilled are required to pledge a lower payoff to the depositors. This eases the diversion constraint and allows the bank to lever up more.

   Combining the above predictions we have that,

3. *Bank Leverage and Size are correlated in the cross-section.*

   Gropp and Heider (2010) provide empirical support for this prediction.

4. *Bank becomes larger and better capitalized as the monitoring cost, F, falls. The leverage ratio is unaffected. (From Equations (4.6) and (4.7))*

   As cost of monitoring, F, falls, the optimal size of equity increases, without affecting the leverage ratio (so bank size increases).

   We may think of a decrease in F as an improvement in the bank's monitoring technology. As the monitoring technology improves, banks become larger due to a higher level of equity capital in the bank while holding the leverage ratio constant.

The existing literature (e.g. Holmtrom and Tirole, 1997) suggest that higher levels of equity will lead to better monitoring incentives. We propose that the direction of causality may go in the opposite direction as well.

4.6 Conclusion

We present a model of bank capital structure in which the investors' outside option determines their choice of security in the bank. The skilled agents may divert from the unskilled in a banking coalition. If diversion is large enough, the optimal arrangement entails the following: the skilled investors hold equity capital and the unskilled investors are the depositors (markets for equity and deposits are segmented). Due to their higher outside option, the skilled command a skill premium and equity is therefore more expensive than deposit.

The bank's leverage ratio is determined by the diversion constraint. Further, the level of equity is determined by the trade-off faced by the skilled agents - economies of scale (benefit) versus monitoring against diversion (cost). These two results are combined to derive the optimal levels of debt and equity from which we get the optimal size of bank.

Bibliography

Admati, A., P. DeMarzo, M. Hellwig, and P. Pfleiderer (2011): "Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Requirements: Why Bank Capital is Not Expensive," *Working Paper*.

Allen, F., and E. Carletti (2013): "Deposits and Bank Capital Sructure," *Working Paper*.

Allen, F., E. Carletti, and R. Marquez (2011): "Credit Market Competition and Capital Regulation," *Review of Financial Studies*, 24(4).

Calomiris, C., and C. Kahn (1991): "The Role of Demandable Debt in Structuring Optimal Banking Arrangements," *American Economic Review*, 81(3), 497–513.

Cerasi, V., and S. Daltung (2000): "The Optimal Size of a Bank: Costs and Benefits of Diversification," *European Economic Review*, 44, 1701–1726.

Christelis, D., T. Jappelli, and M. Padula (2010): "Cognitive Abilities and Portfolio Choice," *European Economic Review*, 54, 18–39.

Cole, S., A. Paulson, and G. K. Shastry (201): "Smart Money? The Effect of Education on Financial Outcomes," *Review of Financial Studies*.

Diamond, D. (1984): "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51, 393–414.

——— (2004): "Presidntial Address, Committing to Commit: Short-term Debt when Enforcement is Costly," *Journal of Finance*, 59(4), 1447–1479.

Diamond, D., and R. Rajan (2000): "A Theory of Bank Capital," *Journal of Finance*, 55, 2431–2465.

——— (2001): "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking," *Journal of Political Economy*, 109, 287–327.

Grinblatt, M., M. Keloharju, and J. Linnainmaa (2011): "IQ and Stock Market Participation," *Journal of Finance*, 66(6), 2121–2164.

Gropp, R., and F. Heider (2012): "The Determinants of Bank Capital Structure," *Review of Finance*, 14, 587–622.

Guiso, L., M. Haliassos, and T. Jappelli (2002): in *Household Portfolios*Cambridge, MA. MIT Press.

Guiso, L., and P. Sodini (2013): "Household Finance: an Emerging Field," in *Handbook of Economics and Finance*, ed. by G. Constantinides, M. Harrris, and R. Stulz, pp. 1397–1532, Amsterdam, North Holland.

Hart, O., and J. Moore (1995): "Debt and Seniority: An Analysis of the Role of Hard Claims in Constraining Management," *American Economic Review*, 85(3), 567–585.

Holmstrom, B., and J. Tirole (1997): "Financial Intermediation, Loanable funds and the real Sector," *Quarterly Journal of Economics*, 112(3).

Mehran, H., and A. Thakor (2011): "Bank Capital and Value in the Cross-Section," *Review of Financial Studies*, 24(4).

Morrison, A., and L. White (2005): "Crisis and Capital Requirements in Banking," *American Economic Review*, 95(5), 1548–1572.

Myers, S., and N. Majluf (1984): "Corporate Financing and Investment Decisions when Firms have Information that Investors do not have," *Journal of Financial Economics*, 13, 187–221.

van Rooij, M., A. Lusardi, and R. Allessie (2011): "Financial Literacy and Stock Market Participation," *Journal of Financial Economics*, 101, 449–472.