Credit Rationing and Relationship Banking

Nicolas Taillet and Michael Tröge

ESCP Europe

Abstract

The paper develops a new theory for why some banks implement loan policies that exclude profitable but risky firms. We assume that firms seeking credit will try to evaluate the performance of a bank's existing client portfolio. In case many of the existing clients have entered financial distress, new clients will suspect the bank to be at the origin of the firms' bad performance. In order to signal that they will not endanger the financial health of their borrowers, banks might need to ration riskier, yet profitable, firms.

Keywords: Bank Reputation, Credit Rationing, Relationship Banking

JEL-Classification numbers: G21, D45

1. Introduction

It has often been alleged that banks implement credit policies that exclude risky but (on average) profitable borrowers.[1] Direct evidence on this type of credit rationing[2] is difficult to come by, as it is impossible to verify whether a loan that has not been awarded would have been profitable. However, indirect evidence comes from the emergence of different types of non-bank lenders such as finance companies (Carey et al., 1998) and more recently private lenders (Arena, 2011, Nini, 2008). These lenders seem to be able to thrive on the riskier part of the loan market that seems to be neglected by traditional banks.

In this paper we argue that credit rationing might be related to relationship banking. In a nutshell, if relationship banks want to acquire a reputation of supporting struggling borrowers, lending to risky companies can damage this reputation. Potential borrowers will observe the high default rate among the banks' clients and attribute this to a lack of support from their banks. This in turn will affect the profitability of the bank's relationship with its existing and potential clients. Hence, there exist reputational spillovers that prevent banks which want to position themselves as trust-based relationship lenders from being at the same time active in the market for highly risky loans.

The paper develops this argument using a signalling model and shows that in case of relationship lending based on reputation, banks can improve their overall profitability by implementing a restrictive lending policy, i.e. by not providing loans to risky clients, even if analysed individually the relationship with these risky clients would be profitable. This behaviour of refusing risky clients is perceived positively by the market, which then identifies the bank as a relationship bank.

The basic rationale driving our results is very simple: some of a bank's actions, such as the early withdrawal of loans, may be appropriate for a firm in real difficulties but can also be interpreted as a lack of commitment and competence on behalf of the bank. If firms are not able to distinguish between a bank withdrawing a loan from a fundamentally

---

[1] Some references even from newspapers our other

[2] Note that here lenders discriminate which is different from simple rationing. Stiglitz and Weiss (1987) refer to this type of rationing as "criterion b rationing".

healthy firm and a bank saving its credits in a failed company, they may misinterpret a justified action as lack of competence or commitment. Firms will want to avoid these types of banks or only choose these banks in case of very attractive interest rates. Banks will therefore prefer not to finance too many risky firms. Note that, whereas the bank is not risk averse in a strictly microeconomic economic sense of having a concave utility function, it is risk adverse in a common language sense, as it refuses to finance risky but profitable firms.

Our paper is complementary to two segments of the literature. The paper is obvioulsy contributes to the large and old literature on credit rationing going back at least to Stiglitz and Weiss (1981), Thakor (1987), Boot and Thakor (1984, 1994). However the explanation for this rationing behaviour is related to the concept of relationship banking which has been analysed is an equally large but separate literature going back to papers such as Chan, Greenbaum and Thakor, (1986), Rajan, (1992) or Rajan and Peterson (1994).

In the seminal Stiglitz and Weiss (1981) paper, credit rationing ultimately derives from the bank's inability to raise a sufficient supply of funds given the maximum return that can be extracted from a group of borrowers. However, Stiglitz and Weiss' approach has been challenged on empirical (Berger and Udell, 1992) as well as theoretical grounds (Riley, 1987, Arnold and Riley, 2009). Unfortunately, while credit rationing is readily acknowledged by bank practitioners, finding hard empirical evidence is a difficult task. In general, it is not possible to decide whether the bank has turned down a lender because at high interest rates, he would not have been profitable (for example because of the increased danger of moral hazard), or because the bank was following a certain credit policy (which lead to not to finance risky but profitable companies). Several empirical observations seem to indicate that even banks with ample funding under-supply credit to risky ventures (Machauer and Weber (1998)) and leave this market segment to other financial agents (Carey et al., 1998). Part of the evidence comes from the interest rates quoted by the banks. The few existing studies on loan rates show that rates do not differ very much among firms of different risk. For example, Machauer and Weber (1998) analyse the relationship between interest rates and the borrowers' bank internal ratings for a large German bank. In their sample, the average interest rate difference between borrowers in the best and the worst of five risk classes is only 1.2%. Given that the best

class is defined as "good or very good creditors", whereas the worst class consists of borrowers which are "very much in danger of default", these interest differences are much smaller than the interest rate spreads on corporate bonds of comparable ratings. This is in line with the results of Petersen and Rajan (1994, 1995) indicating that interest rates depend more on the firms' bargaining power or credit market competition than on risk. In general banks seem to quote rates that are very close to internal prime rates, which indicates that they do not seem to make the interest rate adjustment which would be necessary to profitably lend to risky borrowers. Additional indirect evidence for credit rationing comes from the success of finance companies. Facing the competition of a large incumbent banking sector, these non-bank lenders have succeeded to turn lending to risky companies into a highly profitable business. As demonstrated by Carey et al. (1998) they serve riskier borrowers, which could indicate that these types of firms have been neglected by traditional banks.

Our paper provides an entirely different explanation for why banks might exclude risky but profitable borrowers. As we will discuss in more details, the management and policy implications of our approach are fundamentally different from the insights that can be derived from the traditional explanation of credit rationing behaviours. Indeed, our explanation for rationing is related to reputational mechanisms as presented by Dinc (2000). It derives from the potential positive effects of relationship banking (Petersen and Rajan (1995)) on firms' financing. Rajan (1992, 1994) explains that firms, once they have received credit from a bank, are to some extent captive and highly dependent on the bank's actions. Indeed, there are many concrete ways for banks to behave opportunistically once a credit has been given. For example, banks may try to renegotiate the interest rate or withdraw part of a long-term loan and force the borrower to lend at much higher overdraft rates. In countries with close bank firm relationships many credit contracts contain a materially adverse conditions clause (MAC), which enables the bank to call back credit under conditions which are not precisely specified. Even when this MAC clause does not exist, covenant violations may enable the bank to threaten the withdrawal of a loan or renegotiate the credit contract. These covenant violations are surprisingly frequent. For example, 37% of the firms in the sample analyzed by Chava and Roberts (2008) are in violation of a covenant. Note that increasing interest rate is not the only way to extract rents. Other alleged tactics include the forced selling of expensive services once the client has become captive.

However, in this paper we use another type of opportunistic behavior first proposed by Chemmanur and Fulghieri (1994) and Dinc (2000). In our model, the competitive advantage of a bank with respect to bond finance derives from its ability to distinguish a severe problem which requires liquidation of the firm, from simple covenant violations, which can be solved by restructuring the loan. An opportunistic bank will always withdraw the loan after observing a covenant violation. This does not necessarily lead to an important loss for the bank, but withdrawal will typically lead to financial distress or liquidation of the company. A welfare maximizing bank should not immediately withdraw the loan but should be able to finance the company since the initial project was deemed profitable. The importance of reputation in limiting the opportunistic behaviour of banks has been recognized by several authors. Sharpe (1990) simply assumes that banks acquire a reputation not to exploit informational rents. Similar to our paper, Chemmanur and Fulghieri (1994) focus on the decision of the financier to liquidate a firm in financial distress or renegotiate the credit, whereas Dinc (2000) considers the incentives of banks to observe discretionary loan commitments.

Our paper uses a relatively simple and intuitive model of signalling as first used by Spence (1973). Contrary to the models with a finite number of repetitions in the spirit of Kreps and Wilson (1982) or Milgrom and Roberts (1982) used for example by Chemmanur and Fulghieri (1994), we do not have to assume the existence of a "commitment" player. Moreover, contrary to the simple analysis of infinitely repeated games which have frequently been applied in finance (Sharpe (1990), Dinc (1997)), we do not assume perfect observability or use trigger strategies. The main focus of this paper, however, is not the reputation mechanism itself but its possible drawbacks. The fact that reputation building can lead to suboptimal outcomes might be surprising, but this is not a new finding. Ely and Valimaki (2005) as well as Ely, Fudenberg and Levine (2008) have analysed models with a similar intuition, but with a very different structure.

Our theory of the bank's rationale for credit rationing has a range of interesting consequences. Especially, reputational equilibria will not always be sustainable. In several economic environments, banks may find it useful to apply different lending modes. Specifically, in risky environments, the reputational lending mode might imply the exclusion of too many borrowers. In this case, non-reputation lending to a larger portfolio of clients may be more profitable than maintaining reputation through rationing. Hence,

a key insight of the model is that for explaining credit rationing it may not be sufficient to look at a single bank-firm relationship. In fact, the feasible lending modes depend on the risk profile of the entire economy. A given bank may be willing to finance the same firm in one particular economic environment but not in another.

In this context, we use a similar framework as Winton and Yerramilli (2015), but not in the context of originate to distribute. Indeed, our core assumption is that the loan is illiquid before default, i.e. the banks cannot exit a relationship before the term of the outstanding loan. Still, we keep some of the key features of the model such as the fact that "monitoring lowers the probability of default but does not eliminate it" and that monitoring is socially beneficial. In addition, we distinguish between two bank business models – with and without monitoring – which profitability depends on a) the social gain from monitoring and b) the proportion of low default probability firms vs high profitability firms. In this context, the bank should define a portfolio strategy, rather than focusing on the respective profitability of single borrowers in order to maximize its reputational rent due to the standards it sets. Without the monitoring feature, banks actually behave like any other financial companies, which would contradict the findings of Carey et al. (1998), which showed that finance companies focus on lending to riskier borrowers.

Our paper also has important managerial implications. Banks have in the past invested considerable effort in analysing the risk adjusted profitability of a relationship with individual corporate customers. This paper argues that for establishing a profitable relationship banking strategy it might be necessary to go beyond the analysis of individual clients and to focus on the profitability of an entire portfolio of current and future client relationships. The two approaches are not equivalent. Indeed, maximizing the profitability of each client is not equivalent to maximizing the profitability of a portfolio of clients, if there are negative spillovers from financing some clients on the profitability of other clients. We demonstrate in this paper that bank reputation can lead to exactly this kind of negative spillovers. Banks are concerned that if many firms in their loan portfolio fail, current and future clients will attribute these failures to the bank's action rather than to the firms' business risk. In this case, the bank will not be considered as a relationship bank and will not be able to capture part of the economic surplus generated by the firms.

The main contribution of this article relies in the proof that even in a state where banks can determine borrowers' credit quality, there will be credit rationing due to the very nature of relationship baking mechanism. We think that in countries with extensive relationship banking, credit rationing is a widespread phenomenon with potentially important implications for economic growth.

Going beyond the strict results of the specific model we present, we think that these kinds of models will be essential to convince banks about the specificity and profitability of the business model of relationship banking. The originate to distribute business model has led to a disappearance of borders between bank loans, other financial company loans and bond markets. Yet, this article shows that relationship may be rationally thought as a separate market created by monitoring entities and based on a virtuous sequencing of reputational rents based on observable virtue.

In the next section of this paper we present the model. In Section 3 we develop its implications regarding the existence of potential equilibria. In section 4, we derive the increase in Welfare from an efficient monitoring banking environment depending on the economic environment.

2. The Model

The objective of the model is to demonstrate the possibility of credit rationing as signalling strategy. In other words, excluding some profitable but risky borrowers – hence reducing the number of clients – can increase the overall profitability of a bank because it will enable a bank to convey to lenders that it will support borrowers in case of distress. To keep the model as accessible as possible we rely on a number of existing building blocks from different parts of the banking literature. We demonstrate that these elements can be combined in a model, which leads to the conclusion that in a specific equilibrium of this model, credit rationing can improve the bank's overall profit.

Hence, we do not intend to fully characterize all equilibria for the presented setup. This would lead to more advanced game theoretic questions which are not related to the motivation for this paper. Neither do we want to claim that this is the only way to model credit rationing. However, we think that the overall intuition is very robust and can be formulated in a range of different models.

2.1. Overall model structure

The model presents two types of agents – a long lived bank and short lived firms engaging in lending relationships with the bank. To keep the model structure simple we only examine two repetitions of a relationship lending game, i.e. we assume that two generations of firms successively borrow from the same bank.

Banks are of two types, either they have the ability to monitor ("monitoring bank" (G) in the following) or they do not have this ability ("non-monitoring bank" (B) in the following). Monitoring banks will analyse the firm's situation in case of financial distress and provide financial support if appropriate. This will decrease the default rate in the banks' credit portfolio. We also assume the existence of a competitive fringe of non-relationship lenders that provide the firm with an outside borrowing option.

In addition to monitoring there is also another way for banks to influence the default rate of their portfolio. Bank observe the borrowers' risk and can influence the default rate of the first generation of borrowers by monitoring their activity. In addition, banks choose a credit policy at $t = 0$, i.e. they choose to exclude borrowers exceeding a certain level of risk. During the next period, at $t = 1$, if the company is not successful, non-monitoring banks withdraw the loan while monitoring banks offer a loan workout. Finally, at $t = 2$, payoffs are realized and distributed.

At Stage 2, new entrepreneurs observe the last default rate of a given bank portfolio before deciding whether to consider it a relationship-bank or not. With this information in hand, they play the same lending game as in Stage 1.

The principal objective of the model is to demonstrate that by excluding certain types of profitable firms in Stage 1, the bank can increase the profitability of the equilibrium played in Stage 2. The general idea behind this is that banks can signal their type – monitoring or non-monitoring – during the first stage, to be identified as monitoring banks by the firms during the second stage and to extract a higher share of the surplus via the implementation of a higher interest rate.

2.2. Firm characteristics

Firms live over three periods in time. At $t = 0$, they require a loan normalized to 1. At $t = 1$ the firm can be in one of two possible states, designed by the letters S and F and occurring with probabilities p, and $1 - p$. In state S there is no doubt that the firm will be successful and deliver the output X at $t = 2$. In the state F the firm is in financial distress, but it can be saved if the bank engages in a loan workout. A workout implies that the firm will be producing the success output X with probability p and 0 with probability $1 - p$.

2.3. Bank characteristics

For convenience we will assume that only one bank – the monitoring bank (G) – monitors the firm. It will always monitor the firms in its portfolio as long as it has established a lending relationship with them.

Credit markets are perfectly competitive, but competitors – non-monitoring banks (B) – will only be able to provide non-monitored financing, where companies are liquidated in states F at $t = 1$. Such a behaviour can be explained via the inclusion of a covenant enabling the bank to withdraw the loan if the company is not in state S. Withdrawal implies liquidation of the company with the bank obtaining nothing but saving potential additional restructuring costs.

2.4. Lending relationships

If the firm is eligible for a loan, i.e. if $pX > 1$, the bank decides whether to provide a loan at $t = 0$ or not. If the bank chooses to finance the project, the company accepts financing for a cost R_G from a good bank and for a cost R_B from a bad bank. The good bank spends a fixed cost of m on monitoring. These costs capture the considerable effort the bank has to spend in order to monitor the company as well as potential restructuring costs which may occur during the process in the distressed times. For clarity purposes, the good bank cannot behave opportunistically and not pay m in state F at $t = 1$.

At $t = 1$, the bank can easily observe that the firm is in financial distress, but it does not know for sure whether it will be successful at $t = 2$ or not. At this point ($t = 1$), the monitoring bank engages in a loan work out (*see Figure 1.*), whereas the bad bank withdraws the loan, which leads to the company's default at $t = 1$ (*see Figure 2.*).

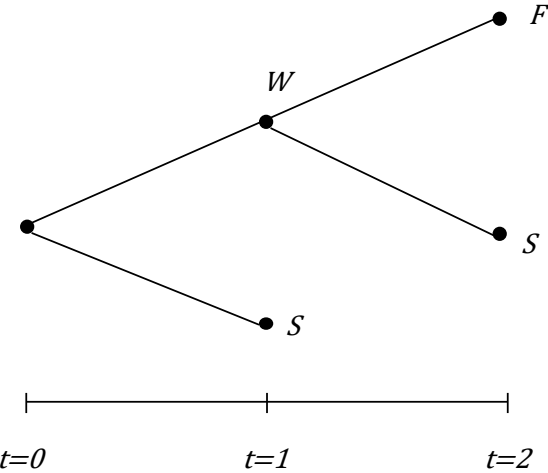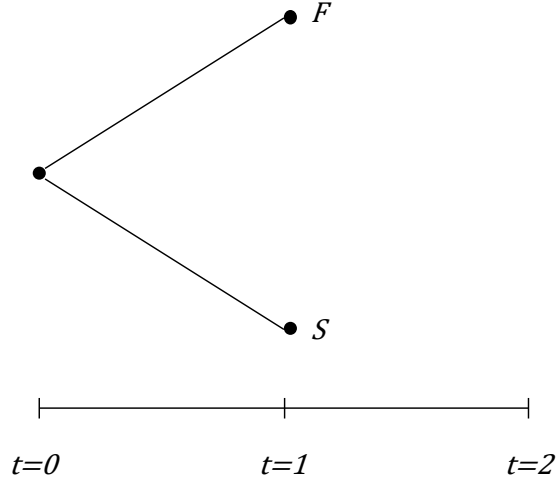*Figure 1. Lending relationship with a monitoring bank*

Figure 2. Lending relationship with a non-monitoring bank

Good and bad banks play the signalling game by excluding the risky companies with higher probabilities of default to decrease the observable default rate of their loan portfolios. A non-monitoring bank hence has the choice between behaving like a bondholder, not discriminating if the borrower is profitable ex ante and trying to mimic the behaviour of good banks to benefit from the same lending conditions.

However, the good bank has an edge because monitoring decreases the probability of default of its portfolio:

- for a non-monitored company with probability of success p the default rate at stage $t = 2$ is $1 - p$
- for a monitored company with probability of success p the default rate at stage $t = 2$ is $1 - (p + (1 - p)p)$ which is lower than $1 - p$

To make our analysis interesting, we assume that monitoring and engaging in a workout is socially efficient, i.e. $(1 - p)pX - m > 0$. This means that banks may have an incentive to help a distressed company, as long as they can capture part of the generated surplus via a higher interest rate, and that society as a whole would benefit from monitoring activities from the bank.

2.5. The signalling game

2.5.1. Description of the signalling mechanism

We assume that the lending game described in the previous paragraphs is repeated two times with the same perpetually lived banks facing a pool of new firms with short term memory in every repetition. New entrants observe the default rate at portfolio level displayed by the bank and draw inferences about the bank's type. However, new firms cannot distinguish between a monitoring bank and a non-monitoring bank displaying the same default rate in stage 1.

Importantly, new firms cannot distinguish whether the firms from the previous set had been in state S or F at $t = 1$, i.e. they do not know whether the observed default rate of the portfolio is due to monitoring or to a bad bank's mimicking a good bank's behaviour via the rationing of a larger part of the firms applying for a loan.

2.5.2 The credit policy

We assume that firms arriving in every period of the reputation game are of two types: θ firms with high probability of success p_H (or low probability of default $1 - p_H$) – thereafter the "good companies", "less risky borrowers" (H) – and $1 - \theta$ firms with lower probability of success p_L – thereafter the "bad companies", "risky borrowers" (L). We have $p_H > p_L$. For clarity purposes both types of firms generate the same potential outcome X in case of success.

In addition, we focus on companies, which are profitable borrowers and directly exclude non-profitable companies, which would not be financed in any case, hence:

$$p_H X > p_L X > 1$$

We also assume that the increased probability of success generated by the monitoring does not decrease the default rate of monitored firms with low probability of success below the default rate of non-monitored firms with high probability of success. Indeed, if it were the case, the good bank could never be copied by the bad bank, hence we assume:

$$p_H > p_L + (1 - p_L)p_L$$

The bank has the possibility to define a credit policy, i.e. to exclude $1 - \theta - \alpha$ ($0 < \alpha < 1 - \theta$) risky borrowers to reduce the observable probability of default of his portfolio. Hence, α represents the number of risky borrowers to be financed by the G-type bank for a given credit policy. The universe of existing profitable borrowers, defined by their probability of default, can be represented as an axis ordering the companies given their respective probability of default (*see Figure 3*):

Figure 3. Universe of profitable companies, default rates
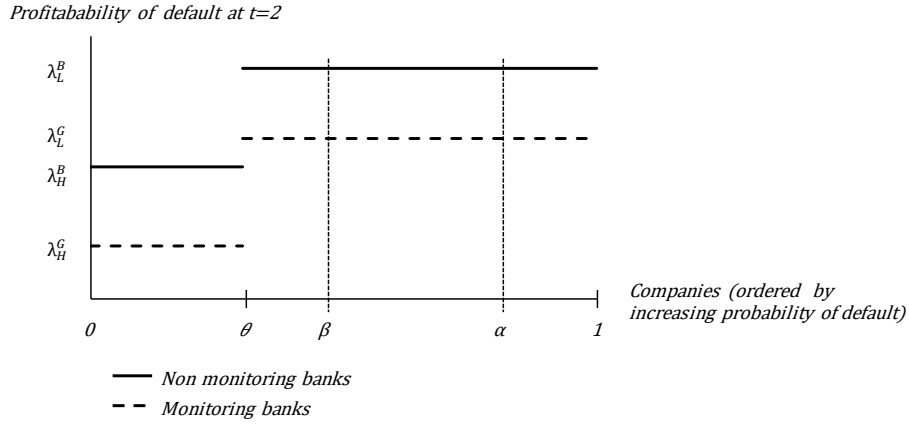
In the chart above, we denote λ_H^G the default rate of a portfolio constituted by only highly profitable companies financed by a monitoring bank (λ_H^B, by a B-bank) and λ_L^G the default rate of a portfolio constituted by only companies with lower profitability financed by a G-bank (λ_L^B, by a B-bank).

By excluding all low probability of success firms, the monitoring banks can prevent any mimicking behaviour from non-monitoring banks. However, it then renounces a large part of the potential profits. The default rate λ_G of a G-bank's portfolio excluding $1 - \theta - \alpha$ risky borrowers (or financing α risky borrowers) can be expressed as follows:

$$\lambda_G = \frac{\theta(1 - (p_H + (1 - p_H)p_H) + \alpha(1 - (p_L + (1 - p_L)p_L)}{\theta + \alpha}$$

To mimic a monitoring bank setting a default rate of λ_G, a non-monitoring bank has to reject $1 - \theta - \beta$ risky borrowers ($0 < \beta < 1 - \theta$). In other words, it only finances β risky borrowers ($\beta < \alpha$) and displays the following portfolio default rate:
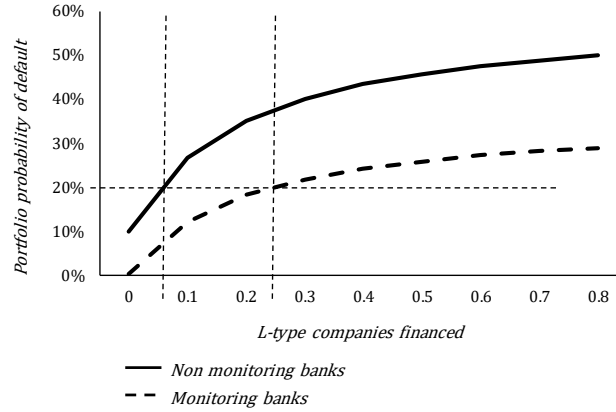
$$\lambda_B = \frac{\theta(1 - p_H) + \beta(1 - p_L)}{\theta + \beta}$$

The credit policy implemented by the G-bank consists in refusing lending to all companies to the right of α (*see Figure 3*) and for a B-bank, in financing only β bad firms.

Below, we show graphically (*see Figure 4.*) the rationing policy to be implemented by a monitoring and a non-monitoring bank, based on the targeted default rate λ. For example, in the chart below, with parameters set randomly, a monitoring bank targeting

a default rate of 20% will be able to finance 0.25 L-type companies, whereas a bad bank will be able to finance only 0.05 L-type borrowers.

Figure 4. Portfolio probability of default

To determine the maximal number α of risky borrowers to be financed while preventing mimicking from B-type banks, we must set α so that λ_G cannot be copied by a non-monitoring bank. Hence, we solve the following equation:

$$\frac{\theta(1 - (p_H + (1 - p_H)p_H) + \alpha(1 - (p_L + (1 - p_L)p_L))}{\theta + \alpha} < \frac{\theta(1 - p_H) + \beta(1 - p_L)}{\theta + \beta}$$

This equation implies:

$$\alpha < \frac{\theta(1 - p_H)p_H}{p_H - p_L - (1 - p_L)p_L} = \alpha_{min}$$

Hence if a monitoring bank decides to finance less than α_{min} risky companies, it can never be copied due to the lower probability of default generated by the monitoring activities.

α_{min} has the following interesting properties:

(i) α_{min} is a decreasing function of p_H, i.e. the greater p_H, the lower the number of riskier companies to be financed by the monitoring bank. Indeed, the higher the initial probability of success of good companies, the smaller the impact of monitoring on the default rate of good companies and the smaller the positive impact of monitoring H-type firms on the probability of default of the portfolio.

(ii) α_{min} is a decreasing function of $\Delta = p_H - p_L - (1 - p_L)p_L$, which represents the delta between the probability of success of a non-monitored H-type firm and the one of a monitored risky firm. The more efficient the bank's monitoring, or the smaller the difference between H- and L-type companies, the closer $p_L + (1 - p_L)p_L$ gets to p_H and the easier it is for a monitoring bank to finance riskier companies without being mimicked.

(iii) α_{min} is an increasing function of θ, i.e. the higher the share of most profitable companies in the sample, the easier it is for good banks to finance less profitable companies without being copied by non-monitoring banks. This also is quite intuitive, since the higher the number of most profitable companies, the lower the default rate in the portfolio and the lower the marginal impact of one additional loan granted to a less profitable firm.

If the bank decides to finance α_{min} risky borrowers, then it will display the following portfolio default rate:

$$\lambda^B_{min} = 1 - p_H$$

Indeed, this is the minimum default rate a non-monitoring bank can display when excluding all risky borrowers to mimic the monitoring bank. By choosing to finance a little less than α_{min} risky borrowers, the monitoring bank cannot be imitated.

We will focus on the cases, where such an α_{min} does exist, i.e. in a universe where the B-type bank has the possibility to mimic the behaviour of the G-type bank. This is true if and if only:

$$\theta \leq \frac{p_H - (p_L + (1 - p_L)p_L)}{p_H + (1 - p_H)p_H - (p_L + (1 - p_L)p_L)} = \theta_{max}$$

Symmetrically, any non-monitoring bank, financing β_{max} risky borrowers and displaying a default rate of its loan portfolio λ_B with:

$$\lambda_B > \theta(1 - (p_H + (1 - p_H)p_H) + (1 - \theta)(1 - (p_L + (1 - p_L)p_L) = \lambda^G_{max}$$

can be identified as a non-monitoring bank (λ_{max}^G being the maximal default rate of a monitoring bank's portfolio). There is a priori no incentive for a B-type bank to ration above β_{max}. If it cannot mimic the monitoring bank, it will finance the all profitable firms.

These conditions restrict the beliefs that the new generation of firms can rationally form about the identity of the bank as a function of the observed default rate.

2.6. Agents' expected payoffs

2.6.1. Lenders' expected payoffs

2.6.1.1. Bad bank participation constraint

A non-monitoring bank withdrawing loans in state F at $t = 1$ and charging an interest factor (i.e. one plus the interest rate) of R_B, with cost of finance normalized to zero, will receive an expected payoff of:

$$R_B p - 1$$

Hence, we can determine the rate R_B for which the bad bank breaks even:

$$R_B^0 = \frac{1}{p}$$

We assume that credit markets are competitive and that the non-monitoring bank will have to offer loans at a price of R_B^0 when identified as such, hence it will make no profits.

2.6.1.2. Outside option financing condition

Just as for non-monitoring banks, the outside option is assumed to be a competitive financial market. It offers financing to profitable companies for an interest factor:

$$R_{out} = \frac{1}{p}$$

2.6.1.3. Monitoring bank participation constraint

Symmetrically, a monitoring bank offering work out loans in state F at $t = 1$ and charging an interest factor of R_G, with cost of finance normalized to zero, will receive an expected payoff of:

$$R_G p + (1 - p)p R_G - m - 1$$

Hence, we can determine the rate R_G^0 for which the good bank breaks even:

$$R_G^0 = \frac{1 + m}{p + (1 - p)p}$$

This means that the interest rate will reflect the efficiency of monitoring (as calculated via R_G^0), i.e. the relation between the monitoring cost and the improvement in ex ante probability of default induced by monitoring measures. The lower R_G^0, the more efficient the monitoring is and the more competitive the good bank is.

2.6.2. Firm's expected payoffs

2.6.2.1. Firm participation constraints

The firm faces the choice between being financed by the outside option at a lower rate or by a bank at higher rate not knowing with certainty that it will be saved in state F at $t = 1$. Indeed, if the bank's type has not been identified by the firm, it has to guess whether it is facing a monitoring bank or not.

The firms estimate the probability ψ ($0 < \psi < 1$), that the bank lending to the firm is a good bank and $1 - \psi$ that the bank does not monitor. This means that – in the absence of signal – the firm will accept to borrow from banks at an interest rate R_{Banks}, which depends on the value of ψ but not on the real type of the bank. Since the firm cannot distinguish between a bank types, it will pay the same rate to any bank offering a loan not

knowing ex ante whether it will be saved or not at $t = 1$. On the other hand, the firm still can borrow from the outside option for an interest factor R_{out}. Hence the firm's expected profit is as follows:

$$\begin{cases} p(X - R_{Banks}) + \psi(1-p)p(X - R_{Banks}), & when\ borrowing\ from\ a\ bank \\ \\ p(X - R_{out}), & when\ borrowing\ from\ the\ market \end{cases}$$

Which translates into:

$$\begin{cases} (p + \psi(1-p)p)(X - R_{Banks}), & when\ borrowing\ from\ a\ bank \\ \\ pX - 1, & when\ borrowing\ from\ the\ market \end{cases}$$

In both cases, we see easily that the firm is profitable as long as:

$$R \leq X$$

That is, as long as the bank does not ask for an interest rate higher than the project's return, the company will accept the loan.

2.6.2.2. Firm incentive to borrow from banks

In equilibrium a bank will be able to quote an interest rate which makes the firm indifferent between a bank and the outside option. Hence the maximum interest a bank is able to quote is obtained by solving:

$$(p + \psi(1-p)p)(X - R_{Banks}) = pX - 1$$

Which leads to the following result:

$$R^*_{Banks} = \frac{1 + \psi(1-p)pX}{p + \psi(1-p)p}$$

Note that this interest rate R^*_{Banks} may not be sufficient for monitoring banks to make non-negative profits.

Proposition 1: the lower the probability for firms to face a monitoring bank, the harder it is for monitoring banks to finance H-type borrowers and hence to signal their own type.

Proof (in Appendix)

Monitoring banks will make positive profits only if:

$$R^*_{Banks} \geq R^0_G$$

We denote the function $f(\psi, p) = R^*_{Banks} - R^0_G$

The function above is growing in ψ but decreasing in p. This means that for a given set up, it is harder to be profitable for H type companies, that is the ones the good banks absolutely need to lend to signal that they are monitoring.

The function above is positive if, ψ is large enough, that is if the companies strongly believe that they are facing a good bank. Indeed, solving $R^*_{Banks} \geq R^0_G$ for ψ, we obtain:

$$\psi \geq \frac{pm - (1-p)p}{(1-p)p[pX - 1 + (1-p)pX - m]}$$

We assume that this is verified, that is that monitoring banks make positive profits and do not have to finance stage 1 losses with stage 2 profits. In the opposite case, where stage 1 profits would be negative, the optimal strategy for a monitoring bank would always be to exclude almost all borrowers except for a few less risky companies to show a low λ at the end of stage 1. Should stage 2 profits be insufficient to cover stage 1 losses, the monitoring banks would not enter the market in the first place.

In a nutshell, uncertainty regarding the type of banks companies are facing not only means decreasing returns for monitoring banks, it also implies that the good banks have a harder time lending to less risky borrowers and hence a harder time sending the signal they will need in the next stages. If firms believe that there are mostly non-

monitoring banks, monitoring banks may not even exist due to lending to the less risky firms not being profitable.

2.7. Impact on social welfare

Excluding any rationing considerations, the social welfare generated by an ex ante profitable profit financed by a loan given the following assumptions amounts to:

- No monitoring bank exist:

$$W = pX - 1$$

- Only monitoring banks exist:

$$W = pX + (1 - p)pX - m - 1$$

Note that in this case and if the monitoring banks are easily identified as such, we have:

$$R^*_{Banks} = \frac{1 + (1 - p)pX}{p + (1 - p)p} = R_G$$

R^*_{Banks} is greater than R^0_G because $(1 - p)qX > m$ by definition, else monitoring would not be socially efficient, which means that if the monitoring bank is identified as such, it will make positive profits.

With such a rate, the monitoring bank lending to a company with initial probability of success p will make a profit of:

$$\pi_G = (1 - p)pX - m$$

The profit is positive and represents the social welfare surplus generated via the monitoring activities.

However, the fact that banks with better signal make higher profits on individual loans once identified does not necessarily imply that the monitoring bank has an incentive

to keep signalling its type. Indeed, the monitoring bank may renounce signalling if rationing means renouncing to too much profits. The next section will analyse in detail under which circumstances monitoring can be sustained in a repeated game.

2.8. Solution approach

The game can be solved in three steps: first, we define the firm's beliefs. Second, we define the strategies that banks can implement, based on their respective types Finally, we analyse the resulting equilibria.

3. Equilibria in a two-stage game

In a two-stage game, the first stage is the period used by the bank to establish a credit policy. The bank knows that if it is identified as a good bank at the end of the first stage, it will be able to capture the whole surplus generated by monitoring during the second stage of the game. However, to do so, it will have to exclude part of the profitable borrowers during the first stage of the game.

3.1. Firm beliefs in a two-stage game

During the first stage, the firm will not be able to distinguish between a good or a bad bank. The firm will choose a bank over the outside financing option (i.e. it will choose to pay the interest rate premium asked by banks) only if the condition characterizing the firm incentive constraint presented above is realized, i.e. if the banks accept to finance the company at an interest rate:

$$R_1 = \frac{1 + \psi(1-p)pX}{p + \psi(1-p)p}$$

3.1. Firms' beliefs

During the second stage, the payoff of new firms will depend on whether the monitoring bank has been identified as such or not. Let us denote by α_1 and β_1 the number of risky borrowers financed respectively by good and bad banks during the first stage.

Below, we provide the rational set of beliefs of companies of the second generation based on the observation of a given strategy during the first stage of the game. We denote β_{mim} the minimum number of L-type companies – resulting in a portfolio default rate of λ_{mim} – for which the B-type bank is indifferent between mimicking the monitoring bank and financing all companies.

The bad bank will not finance less L-type firms than β_{mim} because beyond this point rationing is not as profitable as financing all companies and being identified as a bad bank during the second round. We denote α_{mim}, the number of L type companies financed by the good bank to display a default rate of λ_{mim} at the end of stage 1.

We also denote $P(G|\lambda)$, the probability for a bank to be a good bank based on the probability of default it displays at the end of stage 1.

Figure 5. Firms' beliefs
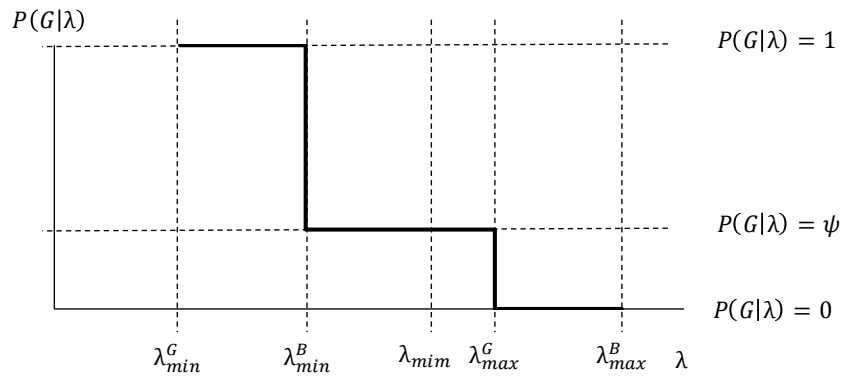
Case 1: The G-type bank finances $\alpha_1 < \alpha_{mim}$ risky borrowers at stage 1, hence displaying λ_{min}^B

The second batch of companies sees the portfolio's default rate at the end of the first stage. The firms identify the good bank as a good bank immediately since they know that no bad banks could have displayed such a low default rate. Hence, the company will accept a loan offer from a monitoring bank if as described above:

$$R_G = \frac{1 + (1-p)pX}{p + (1-p)p}$$

In this case, the firms from the second batch know for sure that they are working with a good bank who will save them in case of distress at $t = 1$.

The companies of the second batch highly value the information from the first generation and are ready to reward a monitoring bank, knowing that they would be saved in case of distress. This generates an incentive for the companies to share the surplus with banks, even in a competitive banking environment.

Case 2: The monitoring bank finances $\alpha_1 \geq \alpha_{mim}$ risky borrowers, the non-monitoring bank finances $\beta_1 < \beta_{max}$ risky borrowers at stage 1

As before, incoming firms at stage 2 see the portfolio's default rate of the bank at the end of stage 1. They cannot be sure whether the bank is monitoring or not since for any portfolio default rate with $\lambda \geq 1 - p_H$, they may face either banks with either types This means, that, similarly to the first stage, the firm will accept to be financed by a bank displaying the same rate as previously:

$$R_1 = \frac{1 + \psi(1-p)pX}{p + \psi(1-p)p}$$

Since ψ is the probability for a company to be facing a monitoring bank in stage 2, the greater it is, the more incentive there is for bad banks to imitate good banks to capture the most surplus possible. If we consider that ψ can be seen as the share of monitoring banks in the banking population, this implies that the greater the share of good banks in the industry, the greater the incentive for free riding incentive.

Case 3: The bad bank finances $\beta_1 \geq \beta_{max}$ risky borrowers at stage 1

In this specific case, where the non-monitoring bank has financed a high number of risky companies during the first stage, and hence where the portfolio default rate

displayed by the bank at the end of the first period is above λ_{max}^G, the firms know that they are facing a non-monitoring bank, hence the bank receives R_{out}.

Note that R_G, R_1 are both greater than the market rate for bad banks. Hence, for the firm, it would be rational to go to the market. This initial consideration shows that without signalling possibilities, non-monitoring banks do not generate additional welfare surplus but only capture part of the surplus generated by the firm due to the activity of monitoring banks.

3.2. Bank strategies

Contrary to firms, banks survive period 1 and grant loans to firms from the first batch as well as firms from the second batch. Hence, there will be some strategic considerations already in period 1 as to what happens in period 2. Since banks know how their behaviour at stage 1 will impact the perception of firms regarding their type at stage 2, they can derive the ex-ante expected profits for both stages depending on the chosen credit policy at stage 1.

The incentives to monitor (or the ability for good banks to generate positive profits) will depend on the profitability of monitoring, which depends on the cost of monitoring, the share of companies with high probability of success, and the perception of the bank's type ψ based on the firms' estimation of the share of good banks in the universe. As mentioned above, p_H, Δ and θ are driving the marginal number of less profitable companies to be financed.

The bank only must choose a strategy for the first stage. Indeed, in stage 2, we assume that the bank finances the whole second batch at R_G, R_1 or R_B depending on the first period strategy. This means that during the second stage, banks will finance all profitable companies since no signal has to be issued for a following period.

Hence, the strategic choice of the monitoring bank is limited to its credit policy at stage 1 as described above. It can choose to clearly signal the fact that it is monitoring by financing $\alpha < \alpha_{mim}$ risky borrowers or to finance more risky borrowers knowing that it will not be identified as a monitoring bank by firms entering the market at stage 2. Symmetrically, the same applies to non-monitoring banks, which will have to choose

between excluding $1 - \beta_1 - \theta$ with $\beta_1 < \beta_{max}$ and being identified as a bad bank. If it is identified as a bad bank, it will make no profits in stage 2.

Below, we provide the overall profit functions of the banks over the two stages, based on their credit policy selected at stage 1.

The profit of the bank over two periods is the sum of the expected profits at stage 1 and of the net present value (NPV) of the expected profits at stage 2, discounted at the discount factor δ.

For clarity purposes, we denote the credit policies as follows in the next paragraphs:

- For the good bank:
  - *sig* means signalling its type at stage 1 by financing $\alpha_{sig} < \alpha_{mim}$ risky borrowers
  - *nosig* means not signalling its type at stage 1, hence financing $\alpha_{nosig} \geq \alpha_{mim}$ risky borrowers
- For the bad bank:
  - *mim* means mimicking the good bank's behaviour and hence not signalling its type at stage 1 by financing $\beta_{pool} < \beta_{max}$ risky borrowers
  - *nomim* means signalling its type at stage 1, hence financing $1 - \theta$ risky borrowers. Indeed, since there are no negative profits, if the bad bank chooses not to mimic, it is always optimal for it to finance all $(1 - \theta)$ risky companies.

3.2.1. Monitoring bank strategies

Strategy 1: The good bank finances slightly less than α_{mim} (denoted α_{sig}) risky borrowers at stage 1 (sig)

$$\Pi_{sig} = \theta(p_H R_1^H + (1 - p_H)p_H R_1^H - m - 1) + \alpha_{sig}(p_L R_1^L + (1 - p_L)p_L R_1^L - m - 1)$$
$$+ \delta[\theta(p_H R_G^H + (1 - p_H)p_H R_G^H - m - 1) + (1 - \theta)(p_L R_G^L + (1 - p_L)p_L R_G^L$$
$$- m - 1)]$$

During stage 1, the monitoring bank receives the same interest rate as any bank R_1^H for less risky companies and R_1^L for riskier companies. However, since its type is

thereafter identified at stage 2, it will receive the interest rates R_G^H and R_G^L respectively from highly profitable companies and less profitable companies.

Strategy 2: The monitoring bank finances $1 - \theta \geq \alpha_{mim}$ risky borrowers at stage 1 ($nosig$)

$$\Pi_{nosig} = \theta(p_H R_1^H + (1 - p_H)p_H R_1^H - m - 1) + (1 - \theta)(p_L R_1^L + (1 - p_L)p_L R_1^L - m - 1) + \delta[\theta(p_H R_1^H + (1 - p_H)p_H R_1^H - m - 1) + (1 - \theta)(p_L R_1^L + (1 - p_L)p_L R_1^L - m - 1)]$$

Just as in case 1 above, during stage 1, the bank receives the same interest rate as any bank R_1^H for less risky companies and R_1^L for riskier companies. However, in this case, it cannot be identified with certainty as a monitoring bank at stage 2, and hence it will once again receive the interest rates R_1^H and R_1^L.

The bank will be incentivized to signal itself as a monitoring bank if and if only the expected profits from strategy 1 exceed the expected profits of strategy 2, that is under the condition that $\Pi_{sig} \geq \Pi_{nosig}$. This means that the benefits from the higher interest rate obtained at stage 2 from being identified as a good bank must compensate for the loss of business caused by the credit policy implemented during the first stage.

3.2.2. Non-monitoring bank strategies

The B-type bank also faces the choice between two strategies:

Strategy 1: The non-monitoring bank finances less than β_{max} (denoted β_{pool}) risky borrowers at stage 1 (mim), to produce a λ similar to the one of a monitoring bank at stage 1:

$$\Pi_{mim} = \theta(p_H R_1^H - 1) + \beta_{max}(p_L R_1^L - 1) + \delta[\theta(p_H R_1^H - 1) + (1 - \theta)(p_L R_1^L - 1)]$$

Note that even though the bad bank may choose to finance β_{pool} ($\beta_{mim} < \beta_{pool} \leq \beta_{max}$) risky borrowers, this would not make sense because it would generate the same

perception as financing β_{max} risky companies at stage 2 and would imply renouncing positive profits at stage 1, hence it will finance β_{max} L-type companies.

During stage 1, the bank receives the same interest rates R_1^H and R_1^L as any bank. Thereafter, it cannot be identified as a non-monitoring bank at stage 2, it will receive the same interest rate R_1^H and R_1^L as during the first stage.

Strategy 2: The non-monitoring bank finances all risky borrowers at stage 1 (*nomim*), and hence cannot mimic the good bank at stage 1:

$$\Pi_{nomim} = \theta(p_H R_1^H - 1) + (1 - \theta)(p_L R_1^L - 1)$$

During stage 1, the bank receives the same interest rates R_1^H and R_1^L as any bank. Thereafter, it is identified as a non-monitoring bank at stage 2 with certainty, it will receive the market interest rate and generate no profits at stage 2.

The bank will mimic the monitoring bank's behaviour as long as:

$$\Pi_{mimic} \geq \Pi_{nomimic}$$

Let us denote, $\lambda_B(\beta_{mim})$ the portfolio default rate of a bad bank financing β_{mim} risky borrowers. The existence of such a β_{mim} depends on the model exogenous parameters, ψ, θ, δ.

Especially, if there is a high preference for immediate results, δ will be close to 0, and hence the incentive of the bad bank to mimic will be lower. The perception of the firms of the economic environment is also important: if the firms know there are numerous good banks in the market, ψ will be high and it may entice monitoring banks to mimic or to extract the maximum surplus during stage 1 depending on δ. Finally, θ has a negative impact on β_{mim}, since the greater the share of good companies in the universe, the harder it is for the bad bank to copy the good one.

If β_{mim} exceeds $1 - \theta$, then it means that the bad bank never has any interest in mimicking.
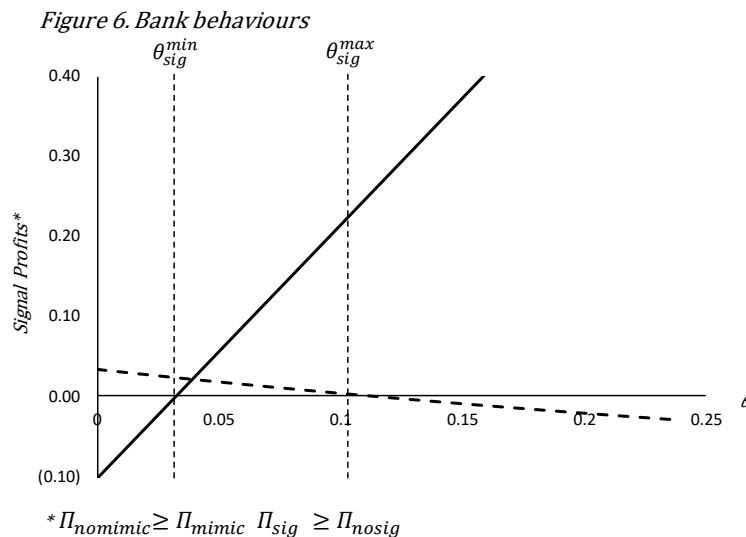
3.3. Model equilibria

Proposition 2: Under control of the assumptions set above, if $\Pi_{nomimic} \geq \Pi_{mimic}$ and $\Pi_{sig} \geq \Pi_{nosig}$, then both type of banks signal their types in the first round and we reach a separating equilibrium, where G-type banks ration credit and B-type banks act like bondholders, not rationing credit.

Proof: See Appendix

This can only happen if θ is large enough ($\theta \geq \theta_{sig}^{min}$) for monitoring banks to be incentivized to signal their type while not being too high ($\theta \leq \theta_{sig}^{max}$) as to encourage non-monitoring to mimic monitoring banks.

This results in second stage companies knowing the type of banks they are facing and hence granting the surplus generated by the monitoring activities of monitoring banks to monitoring banks only. In this case, there is no free riding.

To illustrate this, we draw the decision profit functions of monitoring and non-monitoring banks (*see Figure 6*), and observe where they cross the X-axis, that is where banks have incentive to change behaviour. Where both curves are above 0, the is a separating equilibrium and hence no free riding (when $\theta_{sig}^{min} \leq \theta \leq \theta_{sig}^{max}$). If $\theta \leq \theta_{sig}^{min}$ then


Figure 6. Bank behaviours

$^*\Pi_{nomimic} \geq \Pi_{mimic}$ $\Pi_{sig} \geq \Pi_{nosig}$

Proposition 3: The larger the population of G-type banks in the universe, the harder it is to generate a separating equilibrium.

Proof: see in Appendix

A separating equilibrium only exists if ψ is small enough because as mentioned before, the greater ψ, the greater the incentive for monitoring banks not to signal their type (*see Figure 7.*). In the graph below, we show that with the parameters used above and θ fixed so that $\theta_{sig}^{min} \leq \theta \leq \theta_{sig}^{max}$, only changing ψ, there may or may not exist a separating equilibrium. Indeed, above a certain threshold, monitoring banks stop sending their signal about their type. In this case, i.e. if the share of G-type banks in the universe is very large, the firms at stage 2 know that they are facing a large pool of monitoring banks but they do not identify as such because it would be at the expense of a large rationing on their part.

Figure 7. Impact of ψ on bank behaviour
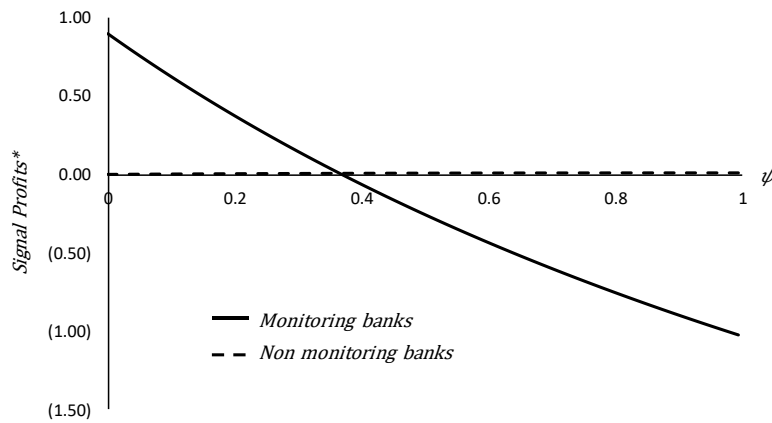
$^*\Pi_{nomimic} \geq \Pi_{mimic}$ $\Pi_{sig} \geq \Pi_{nosig}$

This is interesting because this shows that monitoring banks will only signal their type if the credit quality of the population is rather high overall but if it is too high, then non-monitoring banks will be incentivized to ration and use mimicking strategies, which will grant them a higher share of the surplus than what they should receive since they act as bondholders.

Also, note the importance of the overall share of monitoring banks in the universe. The higher it is, the lesser the incentive for a monitoring bank to signal its type and the greater the incentive for free riding.

The separating equilibrium is the only state where beliefs from companies are perfectly in line with the actual state of things. Indeed, when monitoring banks ration credit, they reach a level of probability of default that the non-monitoring banks cannot match. Yet if the B-type banks do mimic G-type banks, in the belief system as displayed

above, they can still extract part of the surplus via credit rationing. If it were not the case, and for example we displayed a flexible belief system which would change depending not only on the bank's displayed default rate but on all bank's displayed default rates, then there could not be any stable equilibria.

In the case where we make the assumption that monitoring banks would, by showing their types, prevent every non-monitoring bank from being considered as a potential monitoring bank, then we could remove the constraint $\theta \leq \theta_{sig}^{max}$ and there would be a (forced) separating equilibrium, where the B-type banks would have no choice but to play the signalling strategy and not to ration even though.

Proposition 4: Under control of the assumptions set above, if $\Pi_{nomimic} \leq \Pi_{mimic}$ and $\Pi_{sig} \leq \Pi_{nosig}$, then both type of banks do not signal their types in the first round and we reach a pooling equilibrium, where G-type banks do not ration and B-type Banks do ration credit and display similar default rates.

Figure 8. Bank behaviours
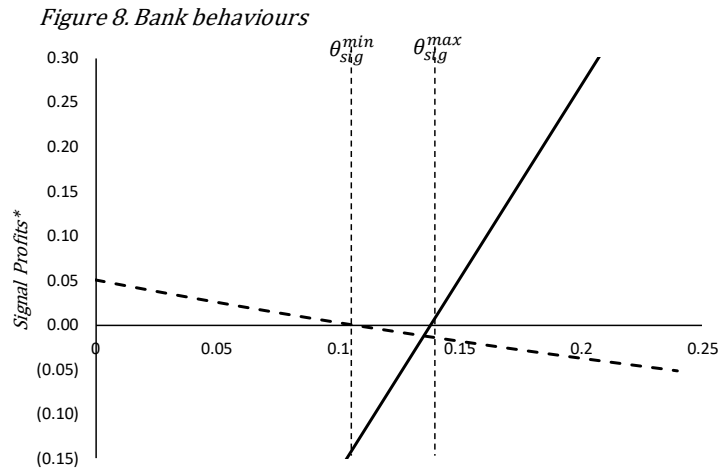
$*\Pi_{nomimic} \geq \Pi_{mimic} \quad \Pi_{sig} \geq \Pi_{nosig}$

As opposed to the separating equilibrium, a pooling equilibrium will happen only in the case where ψ is big enough (for B-type banks to have an incentive to mimic G-banks behaviour) but θ is small enough so that G-type banks do not ration credit.

This is interesting because it shows that if there are too many monitoring banks, they lose the incentive to signal themselves as monitoring banks and hence the non-monitoring banks have a greater incentive to free riding. When a pooling equilibrium is reached, firms are completely uninformed about the type of banks they face and will treat

all banks in the same way at stage 2. However, since such cases appear above a certain ψ, this means that such an equilibrium cannot exist if there is only a small number of monitoring banks in the universe.

Proposition 5: Time urgency (i.e. a low δ) deters banks from rationing. The higher the uncertainty about the future, the less likely it is that monitoring banks signal their type, it also deters non-monitoring banks to mimic.

Proof:

Let us name $f = \Pi_{nomimic} - \Pi_{mimic}$ *and* $g = \Pi_{sig} - \Pi_{nosig}$

$$\frac{\partial f}{\partial \delta} = -[\theta(p_H R_{1H} - 1) + (1 - \theta)(p_L R_{1L} - 1)]$$

$$\frac{\partial f}{\partial \delta} \leq 0$$

$$\frac{\partial g}{\partial \delta} = \theta((p_H + p_H(1 - p_H))(R_{GH} - R_{1H}) + (1 - \theta)((p_L + p_L(1 - p_L))(R_{GL} - R_{1L})$$

$$\frac{\partial g}{\partial \delta} \geq 0$$

This is straightforward, if the banks need liquidity, they will not ration and focus on obtaining the greater profits at stage 1.

The results shown above lead to the definition of each potential equilibria given the optimal reaction function of each bank type. We sum it up in the table below:

If ψ is small enough so that a separating equilibrium may exist depending on θ:

Table 1. Characterisation of equilibria with ψ small

	$\theta < \theta_{sig}^{min}$	$\theta_{sig}^{min} \leq \theta \leq \theta_{sig}^{max}$	$\theta < \theta_{sig}^{max}$
	G-bank		
B-Bank	(No rationing; No rationing)	(Rationing; No rationing)	(Rationing; Rationing)
Equilibrium	Welfare optimal equilibrium	Separating equilibrium	Rationing equilibrium

If ψ is high enough so that a pooling equilibrium may exist depending on θ:

Table 2. Characterisation of equilibria with ψ high

	$\theta < \theta_{sig}^{min}$	$\theta_{sig}^{min} \leq \theta \leq \theta_{sig}^{max}$	$\theta < \theta_{sig}^{max}$
	G-bank		
B-Bank	(No rationing; No rationing)	(No rationing; Rationing)	(Rationing; Rationing)
Equilibrium	Welfare optimal equilibrium	Pooling equilibrium	Rationing equilibrium

As mentioned above, the strategic choices of the banks will be greatly dependent on the model parameters. Still our model shows clearly that it may be rational for a bank to ration credit to send a signal regarding its type. However, it also shows that free riding may convince banks having the choice between monitoring or not, not to monitor in order to extract the maximal surplus generated by the firm, which leads to suboptimal situations.

4. Impact on Social Welfare

The welfare generated by the economy at stage one of a two-stage game can be defined as follows:

$$W_{eco}^1 = \psi\theta[(p_H + p_H(1-p_H))X - m - 1] + (1-\psi)\theta(p_H X - 1)$$
$$+ \begin{cases} \psi(1-\theta)[(p_L + p_L(1-p_L))X - m - 1] \; if \; nosig \\ \psi\alpha[(p_L + p_L(1-p_L))X - m - 1] \; if \; sig \end{cases}$$
$$+ \begin{cases} (1-\psi)(1-\theta)(p_L X - 1) \; if \; nomim \\ (1-\psi)\beta(p_L X - 1) \; if \; mim \end{cases}$$

It is straightforward to see that it is higher when there is no rationing, i.e. when the monitoring banks do not signal their type and when the non-monitoring banks do not try to mimic monitoring banks.

Now let us consider the welfare generated by the economy at the second stage of the game:

$$W_{eco}^2 = \delta[\psi\theta[(p_H + p_H(1 - p_H))X - m - 1] + (1 - \psi)\theta(p_H X - 1) + \psi(1 - \theta)[(p_L + p_L(1 - p_L))X - m - 1] + (1 - \psi)(1 - \theta)(p_L X - 1)]$$

It is completely independent from the banks' behaviour in stage 1. Indeed, only the distribution of the surplus will differ. If monitoring banks have issued a signal at stage 1 then they capture a higher share of the surplus and if non-monitoring banks have rationed in stage 1 then they are capturing a higher share of the surplus than they do when not mimicking. Only stage 1 matters as far as surplus maximization is concerned.

We have shown that rationing may be profit maximizing at bank level given certain parameters, however, it is certainly not at the level of the whole economy.

Proposition 6: The higher the number of monitoring banks in the economy the higher the overall welfare and the lower the surplus extracted by the B-type banks.

Proof: See Appendix

The first part of this proposition is consistent with the model initial assumptions, regarding social efficiency of monitoring. The second part shows that even if at an individual level, the B-type banks profit from a higher number of monitoring banks, the increasing number of monitoring banks implies that a smaller share of the surplus is actually captured by non-monitoring banks.

This means that the incentive for banks to choose to act as monitoring banks has the following effects:

- Lesser incentive for monitoring banks to signal their type and hence to ration credit
- Lesser share of the surplus attributed to free-riding non-monitoring banks at the economy level

We have successfully shown that banks have an incentive to signal their type and ration credit under certain conditions. Although banks have an interest not to monitor and to free ride if not sanctioned, i.e. if firms have difficulties defining their beliefs. Even when they do recognize monitoring banks for what they are, it may be done at a social cost, which may destroy a share of the value generated by the monitoring.

As a first step, let us remind that the higher ψ and the lower m, the easier it is for monitoring banks to enter the credit market and to make positive profits. Hence, the implementation of a tax credit for monitoring expenses may be a solution to make monitoring banks more profitable and to motivate them to exclude as few as possible profitable but risky companies.

The second type of actions to be taken are at information level. The more transparent the market, the easier it is to identify monitoring banks. To increase the market transparency, some additional disclosure about the nature of credit losses may be helpful, metrics such as the amount of time the bank had the defaulting Company in the portfolio (which may indicate some kind of commitment). Transparency in the context of contracts may also indicate the willingness of the bank to restructure: for example, the frequency of contracts with restrictive transferability clauses and flexible financial covenants may indicate that a bank focuses on monitoring and therefore accepts to bear further risks. Further statistics regarding the bank's activity in term of acquisition and sales in the secondary loan markets may also be indicators that the bank is trying to realize short term profits rather than build long lasting lending relationships.

Finally, disclosure regarding the average participation of a bank in a loan agreement may indicate willingness to conduct monitoring activities (in a syndicate, more responsibility is given to the agent, the banks have fewer individual capabilities to have positive actions in the company).

Conclusion

The paper builds on existing literature such as Chemmanur and Fulghieri (1994) and Dinc (2000) showing how reputational concerns can prevent a bank from exploiting the bargaining power that comes with close financing relationships. We show that the reputation mechanism through which cooperation by the bank can be maintained may have potential drawbacks. In particular, reputation building by the bank may lead to rationing risky put profitable companies.

Admittedly, the rationing issue which has been the focus of this paper has a straightforward solution. Banks would be able to offer credit to risky as well as safe companies if they implement a reputational separation between the two types of contract. For example, a bank might set up a finance company as a separate subsidiary providing cheap arm's length contracts to risky borrowers, at the same time it could provide relationship banking through its traditional bank branches.

Our model is also interesting in the sense that it can be used to shed a new light on the concept of a bank's "risk capacity". It is clear that a bank with many safe borrowers might be able to add risky borrowers to its portfolio. As long as these borrowers do not substantially increase the total portfolio risk they will not endanger its reputation. Going beyond the issue of rationing, we think that the general view of relationship banking as a reputational equilibrium used in this paper has wide ranging implications for bank management as well as for financial regulation. Authors like Albach (1997) have argued that reputation and mutual trust is key to understand the advantages of bank financing. Obviously, this idea is diametrically opposed to the widespread approach of simply analysing loans as non-traded bonds. The model developed in this paper illustrates that bank finance is fundamentally different from market-based finance and therefore credit management and credit valuation techniques derived from bond trading and based on fixed income mathematics will miss the key characteristics and the value generated by bank loans. A bank not recognizing that its profits in the credit market are ultimately tied to its reputation and that these reputational equilibria are fragile will not earn any rents.

Management strategies such as giving short term incentives to loan officers or frequently changing bank employees will be detrimental for the reputation of a bank. Unfortunately, it might be difficult for a single bank to implement a reputational

equilibrium as maintaining a high reputation might be difficult in a market where most other players have destroyed their reputation. Tirole (1996) shows that once the reputation of a group has been destroyed, for example by a random shock, it may be impossible to rebuild it. Indeed if most banks behave non-cooperatively, firms might not believe that there still exist trustworthy banks. This industry wide destruction of reputation could also help to better explain the real and long running effects of financial crises.

Discussion

In order to keep the focus of the paper on the core idea that relationship lending can be a viable banking strategy. However, several issues remain to be tackled. Among them, the most important may be the fact that we assume certain banks to be exogenously monitoring or non-monitoring banks. However, banks may be willing to switch strategies based on several parameters, including the economic environment, or the share of the surplus they can capture by monitoring or not monitoring. Such behaviours makes it much more difficult for companies to identify a monitoring bank trying to signal its type, which should also lead banks who choose to monitor not to ration credit but which would also deter banks from monitoring in the first place.

Also, one of the main assumptions of the model is the fact that companies can determine the state of the economy (how many H- and L-type companies there are). If the firms do not know but simply compare the portfolio default rates displayed by banks, they would need to compare the bank default rates without any additional information. This would result in lesser efficiency of signalling strategies.

Bibilography

Arnold, L.G. and J. G. Riley (2009), On the Possibility of Credit Rationing in the Stiglitz-Weiss Model, The American Economic Review, 99, (5) 2012-2021

Berger, A. and G. Udell (1992), Some Evidence on the Empirical Significance of Credit Rationing, Journal of Political Economy, 100 (5), 1048-1076.

Carey, M., M. Post and S. A. Sharpe (1998), Does Corporate Lending by Banks and Finance Companies Differ? Evidence on Specialization in Private Debt Contracting, The Journal of Finance , 53, (3), 845-878

Chava, S. and M. R. Roberts (2008), How Does Financing Impact Investment? The Role of Debt Covenants, The Journal of Finance , 63, (5), 2085-2121

Chemmanur, T. J. and P. Fulghieri (1994), Investment Bank Reputation, Information Production and Financial Intermediation, Journal of Finance, 49, 57-79.

Chan, Yuk-Shee, Stuart I. Greenbaum, and Anjan V. Thakor. "Information reusability, competition and bank asset quality." *Journal of Banking & Finance* 10.2 (1986): 243-253.

Dellarocas, C. (2005), Reputation Mechanism Design in Online Trading Environments with Pure Moral Hazard, Information Systems Research, 16 (2) 209--230.

Dinc, S. (2000), Bank reputation, bank commitment, and the effects of competition in credit markets,\ Rev. Financ. Stud, 13, 781-812.

Ely, J. C. and J. Valimaki (2003), Bad Reputation, The Quarterly Journal of Economics 118 (3) 785-814.

Ely, J. C., D. Fudenberg and D. K. Levine (2008), When is reputation bad? Games and Economic Behavior 63 (2) Pages 498--526

Fudenberg, D. and D. K. Levine (1992), Maintaining a Reputation when Strategies are Imperfectly Observed, Review of Economic Studies, 59, 561-579.

Kreps, D.M. and R. Wilson (1982), Reputation and imperfect information, Journal of Economic Theory, 27 (2) , 253--279

Machauer, A. and M. Weber (1998) Bank behavior based on internal credit ratings of borrowers,Journal of Banking \& Finance, 22 (10--11) 1355--1383

Milgrom, P. and J. Roberts (1982) Predation, reputation, and entry deterrence, Journal of Economic Theory, 27 (2) , 280--312

Petersen, M. and R. Rajan (1995), The Effect of Credit Market Competition on Lending Relationships, The Quarterly Journal of Economics, 110 (2), 407-443.

Petersen, M. and R. Rajan (1994), The Benefits of Lending Relationships: Evidence from Small Business Data, Journal of Finance 49, 3-37.

Rajan, G. (1992), Insiders and Outsiders: The Choice Between Informed and Arms-length debt, Journal of Finance, 45, 1367-1400.

Rajan, G. (1995), Covenants and Collateral as Incentives to Monitor, Journal of Finance, 50, 1115-1147.

Riley, John G. (1987), "Credit Rationing: A Further Remark." American Economic Review, 77, 224-27.

Sharpe, S. (1990), Asymmetric Information, 1069-1087.

Stiglitz, J. E. and A. Weiss (1981), Credit Rationing in Markets with Imperfect Information, The American Economic Review, 71 (3) 393-410.

Tirole, J. (1996), A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality), Review of Economic Studies, 63, 1-22.

Winton, A. and V., Yerramilli (2011) Lender Moral Hazard and Reputation in Originate-to-Distribute Markets . Available at SSRN:http://ssrn.com/abstract=1986956