DECONSTRUCTING RELATIONSHIP BANKING

XAVIER FREIXAS
Universitat Pompeu Fabra

During the last decade the concept of relationship banking has been forged in the theory of banking to reflect the sharing of private information between a bank and its clients and the benefits of a continuing relationship. A number of theoretical contributions have examined the implications of relationship banking on the banking industry market structure, but their results are sometimes contradictory. This paper constitutes an analytical survey that examines relationship banking by means of a simple basic model and studies the implications of relationship banking on the pricing of loans, as well as its effect on the degree of competition in the banking industry.

Keywords: Relationship banking, competition, market structure.

(JEL G21)

1. Introduction

The last decade has witnessed the emergence of the concept of relationship banking. This has provided a major breakthrough in banking theory. Yet, relationship banking applies to a variety of situations and has different meanings in different contexts. Given the importance of this new paradigm, it is interesting to clarify exactly what it is that it brings in and how robust the results obtained are. This is the main motivation of this paper, which we view as an “analytical survey”, where a model is developed so as to analyze in a clear and orderly way the main results obtained in the literature. This, of course, is done at a cost: using a unique model to explore widely different results is like using a Procust’s bed, and some contributions had to be simplified in order to fit our framework.

I would like to thank Fabrizio López-Gallo for research assistance and to two anonymous referees. The author is also grateful to José Jorge for his comments on previous versions of the manuscript. Financial support from MCYT grant no. BEC2002-00429 is gratefully acknowledged.
The term relationship banking is not rigorously defined in the literature. We choose to use this term as referring to the investment in providing financial services that will allow to repeatedly deal with the same customer. Whether the customer has to invest or not in the relationship depends on each specific model. Typically, the standard investment is the one made by the bank in obtaining borrower-specific information. For this reason relationship banking has become synonymous of relationship lending, a restriction we will make here, since our objective is to survey the results existing in the literature. Yet it seems plausible that the complementary view, that of relationship servicing will prove fruitful as well in the future (see e.g. Loranth 2002). The distinction may also be relevant from the point of view of the regulation of competition, since the use of tying by banks is considered anticompetitive, while relationship lending is not (OCC, 2003).

For relationship banking to emerge two conditions have to be met:

1. The bank in place is able to provide services to its client at more favorable conditions than its competitors and this advantage is maintained across time.

2. Contingent long term contracts are not available or cannot improve upon the non-contingent ones.

The first condition may imply better information gathering on behalf of the lending bank, as Berger (1999) suggests, in which case it is crucial that the information is soft information not available to competing banks. The second condition states that, over time, new information may be obtained by the parties and this would lead them to renegotiate any long term contract (see Fudenberg, Holmstrom and Milgrom 1990). This second condition also implies that higher levels of integration, like the merger of the lending bank and its borrower are impossible. Also excluded is the type of contract suggested by Von Thadden (1995) of using a long term line of credit with a termination clause, so that the lender can terminate lending, but if it chooses to continue, it is bound by a pre-specified rate. Therefore, the relevant framework to study relationship banking is the incomplete contracts one.

The first path-breaking contributions to relationships banking are those of Sharpe (1990) and Rajan (1992). The original Sharpe (1990) model, which yields the foundation for relationship banking, states that a
firm’s main lending bank knows its type in an otherwise adverse selection environment, and that this gives the bank an ex post monopoly of information, while its competitors are uninformed and therefore face a winners course. Although initially unnoticed by Sharpe, the actual solution to the game, recently given by Von Thadden (2001), is characterized by the absence of equilibrium in pure strategies (the solution is an atomless distribution). As a result, this set up makes it particularly difficult to explore relationship banking properties. Both Sharpe and Rajan acknowledge that if monitoring provides a better information to the lending bank, it implies that in a multiperiod setting the bank has ex post a monopoly of information, leading to a hold-up situation. In a competitive world, where banks have to make zero profits, the ex post monopoly of information has an implication on pricing: competition for a new customer drives down prices at the early stage until the initial losses make up for the subsequent profits derived from the ex post monopolistic situation. Thus the effect on the price structure is close to the one obtained in switching costs models (Klemperer, 1987).

The second wave of contributions, come from the empirical side and acknowledges the benefits of relationship banking. First, Hoshi, Kashyap and Sharfstein (1990) using a sample of Japanese firms argue that such relationship lending generates a surplus, as firms have an easier access to credit or faster recovery in periods of financial distress. Second, Slovin et al. (1993) showed that the severance of the relationship created a loss in terms of the firms’ market value, as illustrated by the bankruptcy of a large bank, Continental Illinois. Finally, Degryse and Ongena (2001) show that firms maintaining multiple bank relationships are less profitable than those borrowing from a main bank.

Using a different approach, Houston and James (1999) examine a sample of 250 publicly traded firms and find out that “publicly traded firms that rely on a single bank are significantly more cash flow constrained than firms that maintain multiple bank relationships or have access to public debt markets”.

This is interesting because it seems to contradict the Hoshi et al. (1990) result on relationship banking in Japan, where firms engaged in relationship with the Kereitsu bank have more easy access to funding. Nevertheless the number of banks a firm deals with is an endogenous variable, so that the two types of firms in the Houston and James approach have to differ in some exogenous dimension and this may explain the difference.
In addition to establishing that relationship banking is valuable to the firm, the empirical evidence stresses three important determinants of the firms involved in relationship banking: their age, their size and the type of business they perform. The age and size of a firm is relevant because screening successful projects is harder for the bank the younger and the smaller the firm. The type of business is also important because the more intangible the firms’ assets, the more difficult it is to objectively assess its probability of default and its loss given default (Houston and James, 1999).

Concerning the age of the firm, working with a data set of small Portuguese firms, Farinha and Santos (2000) state that “The likelihood of a firm substituting a single with multiple relationships increases with the duration of the single relationship” implying that once a firm has proven to be successful for certain number of periods it develops multiple banking links to avoid the hold-up problem. Considering also the relationship between age and bank borrowing, Petersen and Rajan (1994) obtain that, “… the fraction of bank borrowing declines from 63 percent for firms aged 10 to 19 years to 52 percent for the oldest firms in our sample. This seems to suggest that firms follow a “pecking order” for borrowing over time, starting with the closest sources (family) and then progressing to more arm’s length sources”. This evidence tends to indicate that the younger the firm, the larger the degree of asymmetric information which leaves young firms with the unique option of single bank relationship. Houston and James (1999) also obtain that “bank dependent firms are smaller and less transparent that firms that borrow from multiple lenders” and that “Only 3% of the firms with a single bank relationship have public debt outstanding”. In their results too, once the firm is large, old, transparent and presumably successful enough, it taps the public debt market or engage in multiple bank relations to avoid the hold-up problem.

Regarding firms size, Petersen and Rajan (1994) found that “The smallest 10 percent of firms who have a bank as their largest single lender secure, on average, 95 percent of their loans (by value) from it. By contrast the largest 10 percent of firms obtain 76 percent of their loans from the bank” and also obtain that “On average the smallest firm tend to have just over one lender while the largest firms have about three lenders”.

A third wave of contributions is of interest, as it examines the effect of the degree of competition on relationship lending.Boot and
Thakor (2000) argue that an increase in competition among banks results in more relationship lending, while an increase in competition in financial markets result in less relationship lending. Yafeh and Yosha (2001) address the same issues obtaining similar results in a different framework. Still, Gehrig (1998a, 1998b) and Dell’Ariccia (2000) obtain instead ambiguous results regarding the effect of competition on relationship banking.

Finally, a fourth research avenue has emerged with Berlin and Mester (1999) contribution showing that, in a relationship banking context, core deposits allow for intertemporal smoothing and in this way do not require firms to liquidate at a loss their productive assets so as to repay bank loans when interest rates are high. This is interesting since it allows to explain a puzzle in the theory of banking: banks with more core deposits, defined as the sum of demand and saving deposits, have higher profits.

Notice that the issue of investing in a relationship has been previously explored in labor economics. In the labor contract, both the firm and the employer invest in two types of training: specific training which has only a value within the firm and general training which increases the human capital of the employee\(^1\). In the present set up, instead, it is only the bank that invests and usually the investment cannot be transferred by the firm to another bank, so the bank investment in acquiring information about the firm parallels the employee investment in specific training. This is a striking difference since in the banking literature it is the principal that makes the investment while in the labor literature it is the agent. The second major difference is that we assume the investment of the bank is not observable by the firm, while usually the employee investment can be observed by the employer, even if it is not verifiable.

Since all these results on relationship banking are obtained in models with strikingly different assumptions, our objective in the present paper will be to check their consistency and robustness by providing a

\(^1\)As argued by Acemoglu and Pischke (1999), the distinction is blurred, and the idea that firms should pay for specific training while employees should finance their own general training is contradicted by empirical evidence. Stompler (2000) recent paper exploits a similar idea that a bank investment in monitoring improves its knowledge about the industry, so that the specific investment has an externality on the overall performance of monitoring.
unified framework and examining under what circumstances they hold true.

Our modeling choice has been to focus on the \textit{interim monitoring} that allows the bank to prevent the firm from investing in inefficient projects. The relationship dimension comes from the assumption that banks initially pay a firm-specific sunk cost of monitoring and that once they have paid it, the updating cost is much lower. Thus, the monitoring cost is lower for the bank that has already a lending relationship with the firm.

Using this different set-up allow us, on the one hand, to confirm two basic results of relationship banking: first, that banks will use lower interests rates in the first periods and higher ones at a later stage; second, that relationship banking protects firms against adverse changes in business conditions as optimal contracts with banks endow the firm with some degree of intertemporal insurance.

On the other hand, using a different set-up allows us to clarify some results, and to uncover the lack of robustness of others. Our model allows us to explore, first, the effect of the firms bargaining power on relationship banking. This is a key issue in the literature, as in an incomplete contract set up contracts might be renegotiated. Second, we explore the effect of relationship banking on the effort levels of the bank and the firm and examine how competition affects the intensity of relationship banking.

In the following section we will examine a basic model of relationship banking, with the ex post monopoly it generates as well as the implication it has on pricing. Section 3 will examine the issue of the firm’s renegotiation power. Section 4 will be devoted to the impact of competition on the investment in relationship banking. In Section 5 we will examine how relationship banking provides some intertemporal smoothing that creates value for the firms. Finally, Section 6 concludes.

2. The basic model of relationship banking

In this section we will build a model where banks and financial markets coexist and where banks are able to invest in relationship banking.

We consider three types of risk neutral agents: firms, banks and investors, and assume zero interest rates, an assumption we will relax
in Section 5 in order to analyze intertemporal smoothing of interest rate shocks. Firms have investment projects and borrow either from banks or from financial markets. Thus we restrict all contracts to be debt contracts. The difference between banks and financial markets is that banks monitor their loans while financial markets provide arm’s length finance. The basic reason for relationship banking is that there is an initial fixed cost in monitoring, so that repeated monitoring by the same bank saves on monitoring costs.

Firms have two investment projects to choose from. The good one yields $G$ with a probability $p_G$ and zero otherwise; the bad one yields $B$ with a probability $p_B$ and zero otherwise. We assume $p_GG > 1 > p_BB$ so that only the $G$ project has positive net present value. We also assume that $B > G$. This is possible only if $p_G > p_B$.

We assume there is a continuum of firms characterized by their observable probability of success $p$, where $p_G \in \left(\frac{1}{2}, 1\right), p_B < \frac{1}{2}$.

The firm’s choice is not observable and therefore cannot be contracted upon. To further simplify the exposition, we initially assume that firms do not invest in the relationship with the bank. As mentioned before, this is standard in the banking literature and contrast with the results derived from labor economics. The absence of such an investment imply that banks compete more fiercely, since any firm is able to easily switch from one bank to the other.

Banks have two roles. On the one hand they have an interim monitoring role, as they are able to impose the choice of the $G$ project at a cost $M$. On the other hand, if the firm is in financial distress they are able to help the firm out either by making an orderly liquidation or by making an additional loan. We assume that this will report a benefit $V$ to the firm (alternatively it could benefit both the bank and the firm, but then it will reduce the equilibrium level of the bank repayments, leading to the same results).

Since the surplus $V$ plays an important role, it is important to justify it rigorously. First, the existence of a surplus $V$ may stem from the fact that it is easier to renegotiate with a unique well informed claimholder, the bank, than with many different ones. Renegotiating with disperse bond holders may be extremely costly. Second, banks may have better

\footnote{According to Longhofer and Santos (2000), the existence of such a role for banks is the main reason why they are usually senior creditors when standard incentives theory suggest they should be junior.}
skills at valuing the firm as a going concern and choosing whether to liquidate it or not. This is so, in particular, because banks have access to soft information about the project while financial markets are deprived from it. Third, banks may exist precisely because of their ability to manage assets in case of default. This justification of the role of banks was put forward recently by Diamond and Rajan (2001): a bank is the institution that is able to extract the higher expected value from an investment project with the exception of the project manager.

Firms have the choice of their source of funds, borrowing from financial markets, by issuing public debt, or from a bank which will monitor them. The intermediate case of borrowing from several banks will be considered equivalent to borrowing from financial markets, as we assume it is not profitable to invest in monitoring for any of the banks involved, since none of them is able to appropriate the benefits of the investment.

In a multiperiod setting, the monitoring cost $M$ will be a fixed sunk cost, allowing to enforce the $G$ project at no cost in the future, provided the firm continues to borrow from the same bank. In this way, maintaining a continued relationship with the same bank avoids duplication of monitoring costs. Competition in this set-up will therefore imply that in each period banks quote a repayment for new loans (i.e. for firms which have not been previously monitored) and a repayment for loan renewals.

The degree of competition in the credit market will be modeled in quite a simplified way, by assuming that banks have an expected excess return $\rho$ on their investment.

Taking excess returns as exogenously given is a strong simplification. Still, for the purpose of this paper it appears as a valid one as it simply reflects the degree of monopolistic competition. Namely, the use of an exogenously given excess returns is justified by the one-to-one correspondence between transportation costs in a model of spatial product differentiation and the level of excess returns.

The very existence of positive excess returns implies that there is a scarce resource: monitoring skills or a capacity constraint to operate as a bank. From the perspective of banking theory, an attractive justification for the capacity constraint is the existence of capital requirement regulation combined with a high cost of raising capital which
makes bank capital a scarce resource. The assumption of a scarce capital (either because of capital requirements or because of the fact that banks have to have incentives to monitor the loans they make, as in Holmstrom and Tirole, 1997) is particularly attractive in the present context, as it allow us to study the effect of changes in the degree of competition in the financial market and on the degree of competition in the banking industry.

We will assume a two period model. For the sake of exposition we will first examine the role of financial markets and banks in a static model, we will next introduce relationship banking and end up extending the analysis to the case where the probability of success is random.

2.1 Market Finance

Financial market contracts require a payment \( R \) from the firm that reflects the probability of success \( p_G \). In a one period setting, this implies that funding from financial markets is feasible if the firm is better off by choosing the \( G \) project. This is the case whenever

\[
p_G(G - R) \geq p_B(B - R)
\]  

where \( R \) stands for the gross market rate

This is equivalent to stating

\[
R \leq \frac{p_GG - p_BB}{p_G - p_B}
\]

In addition, we will impose the assumption \( p_B < \frac{G - 1}{B - 1} \) in order to obtain the coexistence of financial markets and financial intermediaries.

Since the gross market rate \( R \) equals \( \frac{1}{p_G} \), constraint [1] becomes :

\[
\frac{1}{p_G} \leq \frac{p_GG - p_BB}{p_G - p_B}
\]

which can be rewritten as

\[
1 + p_BB \leq \frac{p_B}{p_G} + p_GG \equiv \phi(p_G)
\]

Since for \( p_G > \frac{1}{B} \), \( \phi() \) is increasing and \( \phi(\frac{1}{B}) < 1 + p_BB < \phi(1) \), there exists a threshold \( \frac{1}{p_G} \) such that for \( p_G > \frac{1}{p_G} \), the firm is able to borrow from the financial market, while for \( p_G < \frac{1}{p_G} \) it cannot be funded in this
way because it would implement the B project. In other words, there exist (safer) firms that prefer to be issue bonds and other (riskier) that prefer to obtain a loan.

2.2 Coexistence of direct and bank finance

Consider now the role of banks. Banks offer also debt contracts at a different nominal rate $R_M$, and with the benefits $V$ of efficient renegotiation in case of distress. As a consequence we expect $R_M$ to be larger than $R$ as it offers the additional value $(1 - p_G)V$ to the firms.

There are two reasons why a firm may want to be funded by a bank in spite of the higher cost. The first is simply the banks interim monitoring services: it may be the case that firms below $s_J$ would prefer to borrow from financial markets but because of moral hazard are forced to borrow from banks. We will refer to this situation as the pecking order case, as the first choice is market finance and bank finance is a second best (see Diamond, 1991). The second reason is that the firm values the services of renegotiation in case of bankruptcy, $(1 - p_G)V$.

This we will refer to as the horizontal differentiation case, as firms have a choice of the type of funding they prefer, some preferring market finance and others preferring bank finance.

In equilibrium the financial market rate for a firm of risk characteristics $p_G$ is $R(p_G) = \frac{1}{p_G}$ while the bank loan repayment rate is given by $R_M(p_G) = \frac{1 + p + m}{p_G}$. In the pecking order case, the profit the marginal firm $p_G$ obtains by funding its operation in financial market is larger:

$$p_G(G - R(p_G)) > p_G(G - R_M(p_G)) + (1 - p_G)V \quad \text{Pecking Order [4]}$$

the only impediment to market finance comes from moral hazard. As a consequence, the demand for bank loans stem from the firms that are ejected from financial markets.

In the opposite case, of horizontal differentiation, the opposite inequality holds:

$$p_G(G - R(p_G)) \leq p_G(G - R_M(p_G)) + (1 - p_G)V \quad \text{Horizontal Differentiation [5]}$$

and a marginal firm $p_G$ that could tap the financial market, (i.e. $p_G > p_G$) is indifferent between arm’s length finance and bank loans

for some $p_G^*, p_G^* \geq \overline{p_G}$.

The distinction between the pecking order case and the horizontal differentiation one will be relevant for the subsequent analysis of the effects of competition.

The conditions for horizontal differentiation are easily obtained. Using $p_G R = 1$ and $p_G R_M = 1 + \rho + m$, replacing in [6], yields:

$$p_G^* = 1 - \frac{\rho + m}{V}$$

This implies that for horizontal differentiation to occur we need $1 > p_G^* \geq \overline{p_G}$, or, equivalently, $(1 - \overline{p_G}) V > \rho + m > 0$.

In other words, horizontal differentiation occurs whenever the value to the firm of the bank restructuring services is larger than the sum of the monitoring cost and market power spread.

For larger monitoring costs or bank market power, the opposite inequality $p_G^* < \overline{p_G}$ obtains and the pecking order case occurs; for lower monitoring costs, the demand for market finance dries out. Horizontal differentiation occurs in the competitive case, $\rho = 1$ provided that $V > M$ that is, monitoring costs have to be lower than the firm’s gain in case of renegotiation.

**Figure 1**
Firms expected one period profit under bank and market finance
As illustrated in Figure 1, firms with very bad credit rating $p_G < \frac{1+p+m}{G}$ will not obtain credit; firms with very good credit rating ($p_G > p_G^*$) will prefer to raise funds in the market and firms in the intermediate range ($p_G^* > p_G > \frac{1+p+m}{G}$) will prefer bank loans. This set up, inspired of Bolton and Freixas (2000), is interesting here as we may assume that smaller firms have higher probabilities of default. Thus, the model predicts that larger firms will access the financial market while smaller ones will obtain bank loans.

Equilibrium in the loanable funds market requires setting the expected return on the bond market equal to the riskless interest rate, which is here normalized to zero.

For large values of $\rho$, firms prefer bank lending and choose their source of funds according to pecking order. Any small change in the rate $\rho$, leaves the demand and supply of market finance unchanged. In the opposite case, of horizontal differentiation, the marginal firm $p_G^*(\rho, 1)$ and the demand for loans will be sensitive to small variations of $\rho$. The effect of a decrease of $\rho$ is to increase the market share of loans while the market for direct finance is reduced but some new borrowing firms are financed, as the threshold $\frac{1+p+m}{p_G}$ is lowered.

**Figure 2**

The two period investment/return pattern (The bank cash flow appears in the upper row, the firm cash flow in the lower row)
2.3 Relationship banking

Since by definition relationship banking requires a multiperiod horizon, we proceed in this section to extend the horizon to two periods\(^3\), as illustrated in Figure 2 and to determine the price structure.

We will denote by \( R_{N1}^j \), \( t = 1, 2 \), the repayment bank \( j \) demands from a new firm which has not yet been monitored, and by \( R_{N2}^j \) the repayment it quotes to a firm that has already been monitored.

**Proposition 1** Assume the expected excess return on bank lending is \( \rho \). Then the Bertrand equilibrium structure of repayments is the following: \( R_{N1}^j = \frac{1 + \rho + (1 - \rho) m}{p_G} \) and \( R_{N2}^j = R_{R2}^j = \frac{1 + \rho + m}{p_G} \).

**Proof.** During the last period a new loan implying a cost \( 1 + M \) will be made with a repayment \( R_{N2}^j \) such that \( p_G R_{N2}^j = (1 + \rho) (1 + M) \). As a consequence, this is the maximum rate the incumbent bank is able to extract on a loan renewal. Competition at time 1 when the market yield is \( \rho \) implies that the banks have to obtain a profit equal to \( (1 + \rho) (1 + M + p_G) = p_G (R_{N1} + (1 + \rho + m)) \) yielding \( R_{N1} = \frac{1 + \rho + (1 - \rho) m}{p_G} \). As it is intuitive the per dollar profit on each loan is the same independently of the probability of success \( p_G \).

Proposition 1 states that the bank will use its ex post monopoly power during period 2, and the borrower will face a hold-up situation. Therefore competition among banks drives down rates at time \( t = 1 \). In the competitive case, we would observe a second profit in the last period and a loss in the first one, so that the competitive zero profit condition is satisfied over the two periods. Proposition 1 is unsurprising as we are facing an environment quite similar to the switching costs one. We know that the presence of switching costs implies lower prices in the initial periods and higher prices later on.

As it is intuitive, \( R_{R2} > R_{N1} \) as the hold-up occurs during the second period. Therefore the bank will accept to grant credit provided \( R_{R2} \leq G \), so that the bank will finance any project with \( p_G \geq \bar{p}_G = \frac{1 + \rho + m}{G} \), as in the one period loan case.

\(^3\)In a previous version of this work we considered any finite horizon \( T \), as we thought that this would lead to more competition between banks. To our surprise it did not affect our results at all, so we choose to keep the standard two period framework.
It is easy to show that the hold-up problem of period two is not alleviated when the time horizon spans. The fact that the two banks compete at each period implies that the hold-up situation is postponed until the last period when competition is non existent. For low levels of $p_G$ the bank incurs a loss during the first period which it recovers in the last period. For higher levels of $p_G$ the probability of reaching the last period, where the hold-up is supposed to occur declines, but during each period the bank gets back a fraction of the monitoring expenses.

The existence of a hold up is, of course, related to the fact that the two parties cannot sign a complete contingent contract. As a consequence, the issue of renegotiation is here relevant.

The extension to the case where the firm also invest in the relationship is straightforward. If the firm has to invest $S$ as a sunk cost in order to benefit from a bank loan, then during the last period a new loan will cost the firm $1 + p + m + S$. Consequently, whether the cost is due to an investment of the bank or to an investment of the firm is immaterial, as the firm’s sunk cost $S$ affects the cost of the loan exactly in the same way as a monitoring cost $\frac{S}{1-p}$ on behalf of the bank.

Regarding the effects of market power, we obtain the standard result that in a less competitive framework (or alternatively when the cost of capital is higher), access to funds by risky firms is reduced and that prices increase not only in the second period but also in the first one. Thus our results is in line with the standard results on monopoly pricing and in opposition to Petersen and Rajan (1995).

Petersen and Rajan’s result depends crucially on their assumption of adverse selection on the riskiness of the investment project. This provides them with a Stiglitz-Weiss type of model where, as a consequence, an increase in interest rates leads the best borrowers to drop out and increases the default rate on bank loans. Therefore, if an investment success leads to additional value, it is possible that, for some parameter values, a decrease on interest rates in the first period of an investment project and an increase in the second period leads to a Pareto improvement. By subsidizing young firms and extracting rents from the old ones the rationing problem of Stiglitz and Weiss is alleviated.4

---

4 Petersen and Rajan refer to the Schumpeterian theory of monopoly as an incentive to innovate in the context of lending, a point developed by Mayer (1988) which
Still, in the general case, the hold-up problem, as any monopolistic distortion, may lead to an inefficient allocation. This raises an important regulatory issue: assume that it could be possible to force the monitoring bank to reveal its information on a firm, so that any other bank has a zero cost of monitoring it (instead of the cost $M$). Would this be an efficient regulatory measure? Of course, the banks would still have the same per dollar return of $\rho$, so the difference will only be that they will price in the cost of monitoring in the first period. Imposing such a sharing of monitoring might therefore be pointless if firms have fixed size projects. From a practical point of view, this issue arises in the context of credit bureaus, which exist in some countries but not in others.

Padilla and Pagano (2000) point out that the information provided by credit bureaus does not always correspond to making the type of the firm known. In fact many credit bureaus share default information, which is referred to as “black information” but do not share other characteristics of the firm which would be part of the white information and would give a more complete picture of the firm’s type and its probability of default. When credit bureaus share only black information the ex post monopoly power of the bank on defaulting borrowers increases, as other banks face a higher credit risk. In this case, the role of a credit bureau is to improve the entrepreneur’s effort because in case of default the ex post monopoly effect is even stronger.

3. Relationship banking and bargaining power

We have mentioned that the hold up issue emerged in a context of incomplete contracts. It is therefore important to explore what happens depending on the relative bargaining power of the bank and the firm. The previous section assumed that the firm was stuck and required funding from the bank, so that the bank had all the bargaining power. We will now see that the higher the bargaining power of the firm the lower the degree of relationship banking.

Notice that the fact that the firm cannot bargain upon the terms of the last period payment, $R_{t+1}^b$, is essential to the result. To see this point, assume that the firm has all the bargaining power. Then at time $t = 1$ the firm makes a take it or leave it offer that leaves the bank at its reservation level of a repayment equal to $1 + p\bar{c}$. Since at this stage $M$ they do not take any further. This point is considered later on in this paper as it is relevant only when the level of monitoring is endogenous.
is a sunk cost, the bank will accept this offer. Therefore, anticipation of this situation will lead the bank (and its competitors) to ask for a repayment \( \frac{1 + \rho + m}{p_c} \) for the first loan granted.

As a consequence, two extreme cases can be considered. In the first one, which we will refer to as the “tough bank” case and corresponds to a classical relationship banking framework, the firm cannot renegotiate its last period repayment and therefore will be facing a hold-up situation. In the second one, corresponding to the “lenient bank” case, the firm would renegotiate the last period repayment and, because the bank expects this to occur, monitoring costs will be reflected in the rate of the first loan. In both cases the economies on monitoring costs are used. The difference between the two is that in the tough bank case credit rates are increasing over time, while in the lenient banks one they are decreasing.

Notice that in either case, the benefits of relationship banking accrue to the firm even if it has to pay the cost of the banks market power \( \rho \). More generally, we may consider a continuum of assumptions on bargaining power distribution that will lead to intermediate situations.

There are three reasons why the analysis of the firms bargaining power is essential to the understanding of relationship banking.

First, there is some confusion in the literature between the banks market power, which we measure by \( \rho \), and the firms renegotiation power. Comparative statics may refer to one, the other or both simultaneously depending on the model considered.

Second, it should be noticed that when we take the standard viewpoint of the tough bank case, which Proposition 1 reflects, the bankruptcy of a bank does not have any impact on a firm’s profit. If, instead, we take the opposite point of view and assume that firms have all the renegotiation power, then borrowing firms benefit from continued relationship, and its severance implies a loss in profits. This would explain Slovin et al. (1993) empirical results on Continental Illinois bankruptcy.

Finally, renegotiation power is essential to risky firms lending. If the rate \( \rho \) is given, tough banks will finance more risky ventures, as they will ask for a lower initial repayment. This is exactly Petersen and Rajan (1995) point, but it is important to emphasize that the effect is not on \( \rho \) but on the power to extract future rents.
A firm bargaining power may come from exogenous factors, one of them is its ability to access funding in financial markets.

The empirical results of Petersen and Rajan (1994), Farinha and Santos (2000) and Houston and James (1999) that we have briefly reviewed, seem to suggest that there is a "financial life cycle" for firms where relationship banking is the training school for small firms that will, later on, tap the financial market. Our model allows to represent this life cycle by simply assuming that the probability $p_G$ increases through time. The firm will then borrow from the market once it reaches a probability larger than $p_G^*$ as illustrated by Figure 1. The theoretical implication is that, in this context, the bank will hold-up the firm the period before the firm is able to access the financial market\(^5\). So the interest rates will be decreasing in the first periods to reflect increasing probabilities of success, but not necessarily at the last stage of the banking relationship, where the hold-up rate may make up for decrease in interest rate implied by the improvement in credit risk.

Contrarily to Petersen and Rajan (1995), our model does not allow to obtain credit rationing, defined as the difference between supply and demand for bank loans at the prevailing equilibrium interest rates. Still, it allows us to consider what types of firms will be the first ones to be denied credit because of their poor credit ratings. Interestingly, this will again depend upon the bargaining power of the lender and the borrower. Consider first the case of tough banks. These banks will make its monopoly profit during the last period, and therefore will grant credit first to the firms they have already lent to in the past. On the other hand, for lenient banks there is no hold-up and severing a relationship has no cost, as all the benefits of the relationship are appropriated by the firm.

This may illustrate the striking differences obtained in empirical studies in different countries. In Japan, Hoshi et al. (1990) find that firms belonging to a Keiretsu have a lower probability of being denied credit. In contrast, among large quoted firms in the US, those having a relationship with a unique bank have a higher probability of being credit rationed (Houston and James, 1999). A higher renegotiation power of firms in the U.S. (additional outside opportunities to obtain funds) would predict these empirical findings.

\(^5\)This characteristic, which is straightforward in our model, may lead to conflicts of interest if the bank is also the underwriter in a more elaborated set up.
4. Effort level and competition

It seems natural to assume that both the bank and the firm are able to make investments for the benefit of the project they jointly develop. Still, the firm may not undertake efficient investments for fear of an ex post hold up by the bank, while the bank may be reluctant to provide funds for fear of being deprived of its fare share of the surplus (Mayer, 1988).

The incentives of the two parties to invest in the project depend on market power, and therefore, it is a key issue to know how banks market power affects the bank and the firm investments in the relationship.

4.1 The bank’s effort level

In the previous section we were confronted with a zero sum game and even if the marginal cost of borrowing in the second period becomes excessively large there is no distortion and thus no cost of the hold-up, because the firm’s project has a fixed size. Yet, in the general case, there is a cost of a hold-up, as the different parties engage in activities leading to suboptimal choices of the different strategic variables they control (Hart 1996). This is the reason why we now give the bank the choice of a level of monitoring that could vary continuously as a reaction to changes in the degree of competition and give the firm a level of effort it controls.

We consider that the bank has a continuous choice of the monitoring expenses $M$, and assume that this choice has not only the effect of preventing the adoption of the bad project, but also increases the probability of success. Namely, we will assume that $p_G(M)$ which we will denote by $p_G$ for short, is an increasing concave differentiable function.

The efficient level of monitoring, $M^*$, will be given as the solution to:

$$Max_M [p_G G + (1 - p_G) V - 1] (1 + p_G) - M$$

and the solution will be characterized by the first order solution:

$$\frac{dp_G}{dM} [G + 2p_G G - 1 - 2p_G V] = 1 \quad [8]$$

The banks profit maximizing effort level will be determined as the solution to

$$Max_M p_G R_{N1} + p_G^2 R_{R2} - (1 + p_G) M$$
so that the banks disregard the externality they produce on firms: as the probability \( p_G \) increases, the firms increase their profit \( G - R_{N1} \) and \( G - R_{R2} \) in the good states of nature. On the other hand, the bank also disregards the benefits of restructuring \( V \) that accrue to the firm. This has a counterbalancing effect, since an increase in the probability of success decreases the probability of appropriating \( V \). It seems natural to focus on the cases where \( G - R_{R2} > V \), that is those cases where the firm prefers success to failure. The first order conditions for the market solution is then characterized by:

\[
\frac{dp_G}{dM} \left[ R_{N1} + 2p_G R_{R2} - 1 \right] = 1
\]

Since \( G - R_{R2} > V \), this implies that the market leads to an excessively low level of monitoring, \( \tilde{M} < M^* \), since \( \frac{dp_G}{dM} \) is decreasing.

As an obvious consequence, ceteris paribus, the inefficiency is reduced when the bank appropriates a larger part of \( G \), that is when market power \( \rho \) increases. This is the basis for the following result:

**Proposition 2** More competition in the banking industry (lower \( \rho \)) decreases the monitoring effort in each relationship banking loan.

**Proof.** Using proposition 1, a lower \( \rho \) implies lower repayments \( R_{N1} \) and \( R_{R2} \), and therefore a lower equilibrium level of monitoring effort since \( \frac{dp_G}{dM} \) is decreasing. ■

Proposition 2 raises two issues: first, is it possible that a non-competitive framework might increase the efficiency in bank lending? and second, is it the case that in a non-competitive framework more risky projects are funded?

Regarding the first issue, in our set up the answer is straightforward provided we distinguish two effects: the level of effort and the extent of the credit market. A higher level of effort implies a more efficient relationship. Yet a more extended market for credit also implies higher efficiency as banks are better at restructuring firms.

As monitoring effort increases with \( \rho \), an increase in \( \rho \) increases overall efficiency (provided it does not affect the firm’s effort level, as we are implicitly assuming here). This will result in higher profits for the firm only in the case of a very strong sensitivity of \( p_G \) to the monitoring level \( M \). In a more general set up the level of production is not fixed but depends upon the marginal cost of funds. In this case the positive effect
of market power on monitoring is to be compared with the standard market power distortion on marginal cost pricing.

Our proposition states the standard view on the effect of competition on the level of relationship. Still, there is another side to this issue if one considers the amount of relationship banking, measured by the size of the credit market. Thus instead of considering how deep is the relation this alternative view emphasizes how extended is the access to the credit market. From that perspective competition could focus on the respective size of the direct finance and market for credit. The use of our model allow us to consider these effects as illustrated in Figure 1. A reduction in $\rho$ increases the market for loans as on the one hand it decreases the cut-off point $\frac{(1+\rho)(1+M)}{p_G}$ and on the other hand, as it decreases the cost of a loan, it increases $p_G^e(\rho)$, in the horizontal differentiation case or leaves unaffected $p_G$ in the pecking order case. We have thus established the following result:

**Proposition 3** More competition in the banking industry (lower $\rho$) increases the extent of the credit market in the horizontal differentiation case and leaves it unchanged in the pecking order case.

Combining propositions [2] and [3], it is clear that the overall effect depends upon the effectiveness of monitoring on improving the firms’ probability of being successful. If this effect is reduced, more competition increases efficiency in the horizontal differentiation case, provided that the marginal cost distortion is low. If the effect of monitoring on the firms probability of success is large, then competition will reduce efficiency both in the horizontal differentiation and the pecking order case.

Regarding the second issue, initially raised by Petersen and Rajan (1994), that states that banks with higher market power will tend to finance riskier projects, our model shows that the effect of market power on the financing of risky projects depends upon how market power and monitoring affect the higher credit risk to be financed, which we denote by $p_G^e$ and is characterized by $p_G^e(M(\rho)) = \frac{(1+\rho)(1+M(\rho))}{(1+\rho)(1+M(\rho))}$. The condition for the Petersen and Rajan effect to take place is that an increase in $\rho$ increases funding to the higher credit risks $\frac{dp_G^e}{d\rho} < 0$. In our framework $\frac{dp_G^e}{d\rho} = (1 + M) + (1 + \rho) \frac{dM}{d\rho} < 0$. The effect is therefore the opposite. The fact that the effect of market power on relationship lending is not robust does not come as a surprise. Gehrig (1998a, 1998b) shows that the relationship between competition and screening
is ambiguous and so does Dell’Ariccia (2000) in a different type of model.

Our perspective differs from the Boot and Thakor (2000) one as they assume competition affects the credit market and financial markets asymmetrically, and that the effect on the later is stronger than the one on the former. When competition between banks increases, it affects more the profitability of market finance than the one of relationship loans, and as a consequence, banks rearrange their priorities increasing their relationship loans to the detriment of their transaction loans. The intuition is simply that relationship loans allow for a partial insulation of bank loans from competition, and therefore banks facing more competition will rearrange their portfolios by investing more in relationship loans. Yafeh and Yosha (2001) suggest a similar result.

Finally, we should consider how the firm renegotiation power affects the bank’s investment. Unsurprisingly, an increase in the firms power to renegotiate will decrease the bank’s investment to monitor. This is intuitive, as the bank incentives to monitor stem from the fact that with some probability the bank will hold-up the firm. If, in case of hold-up, the firm is able to renegotiate, this decreases the banks incentives to monitor.

4.2 The firm’s effort level

We now turn to the firms investment in the relationship. Paralleling the strategic choice of an effort level by the bank, the firm may also invest in a non-observable variable \( e \), which we call effort, affecting the probability of success\(^6\). As before, we measure the effort variable \( e \) by its cost after due normalization.

The efficient levels of monitoring and effort are jointly determined by equations [8] and by

\[
\frac{d\rho_G}{de} [G + 2p_GG - 1 - 2p_GV] = 1
\]

The effort level chosen by the firm, \( \hat{e} \), will be characterized as the solution to:

\[
Max_{e} p_G [G - R_{N1}] + \hat{p}_G^2 [G - R_{R2}] + (1 - \hat{p}_G^2)V - e
\]

\(^6\)Notice that if the effort affected only the cash flow \( X \) obtained in case of success, the borrowing firm will always choose the efficient level effort as it would internalize all marginal costs and benefits. This is characteristic of the debt contract.
and therefore the first order conditions will be

\[
\frac{dP_G}{d\epsilon} [G - R_{N1} + 2p_G(G - R_{R2}) - 2p_GV] = 1
\]

The term between brackets is lower than \(G + 2p_GG - 1 - 2p_GV\) because \(R_{N1} + p_GR_{R2} > 1\). As a consequence, except in the case of large substitution effects between \(M\) and \(\hat{\epsilon}\), the bank and firm choices, \(\hat{M}\) and \(\hat{\epsilon}\) are both lower than the optimal ones. As in the symmetric case of monitoring effort, this is due to the fact that a leveraged firm does not appropriate all the benefits of a higher probability of success (as the bank also benefits from it) but it bears all the costs.

As it is natural, a higher market power or a higher probability of a hold-up will lower the effort level\(^7\).

5. Intertemporal smoothing

The issue of intertemporal smoothing has been developed in an elegant model by Allen and Gale (1997) as a justification of financial intermediation. Their argument is that intertemporal smoothing provides insurance for risk averse agents which would otherwise be confronted with random market conditions. Using an overlapping generations model, Allen and Gale show that the allocation obtained through a financial market alone is ex ante inefficient and that therefore, there is room for another type of contract that would help consumers to obtain intertemporal insurance. Still, Allen and Gale fail to demonstrate that banks could provide this type of contract. Berlin and Mester (2000), develop a somewhat similar argument where they focus on firms rather than on consumers and base the demand for insurance on the existence of liquidity costs, these costs being defined as the opportunity cost of the firm’s lack of cash.

We will introduce this insurance dimension within the relationship banking in a simplified way, by assuming that, with a probability \(p\) the riskless interest rate is \(r^L\) and that with the complementary probability, \(1 - p\), it takes the value \(r^H\). For the sake of simplicity, we assume that credit risk is the same in both states of nature.

Without loss of generality we consider a two period time horizon, and we assume that interest rate risk affects only the second period. In this context, a bank contract is given by a triplet \((R^j_{N1}, R^j_{R2}, R^{Hj}_{R2})\) where

\(^7\)See Padilla and Pagano (2000) for the reverse argument, when in case of default the hold-up problem is more severe, this may increase the firm’s effort level.
the superscript $j$ stands for the bank, $L$ and $H$ for the low or high interest rate and the subscripts $N$ and $R$, as before, for new loans and loan renewals. The bank is able to commit to the rates it quotes. Still, since we are referring to loans and not to loan commitments, it is only natural to assume that the bank has the option not to lend at all.

Within the framework of relationship banking, it may be argued that if banks are able to commit, they might as well be able to commit not to hold up the firm, as argued by Von Thadden (1995). To some extent, the hold up issue, which is essential to prove that interest rates will increase during the last period, is immaterial for other aspects of the relationship. It is therefore possible to assume here that the bank is able to commit and that no hold up will occur, in which case relationship banking will be justified by the fact that the cost of operating with the same bank is lower. Alternatively, a hold-up may occur if this gives the right incentives for the bank to monitor the firm with a higher level of effort $M$.

In order to capture the cost for a firm of interest rate volatility, we will assume that the firm has additional assets at time $t = 2$ that it may liquidate at a cost if necessary. Namely, we will assume that when interest rates are high, $r^H$, a successful firm is able to raise an amount $L$ at a cost $C(L)$, with the standard assumptions $C'(L) > 0$, $C''(L) > 0$. In addition, we take $C'(0) = 1$, so that the marginal cost of raising the first unit of liquidity is zero. This set up allows us to have a cost of high interest rates, provided that we have $(1+p)(1+r^H+M) > G$.

We will show that the bank contract will provide interest rate smoothing and, more important, that when the bank cannot commit to lending, smoothing may be limited by the bank’s option to invest in the riskless asset. In this later case, we will show that only the banks with ex post monopoly rents, (i.e. those we have referred to as tough banks) are able to provide some intertemporal insurance.

**Proposition 4** The equilibrium contingent contract when the bank fully commits to lend is characterized by a second period payment

---

8 The reader will notice that the model is much more general and requires only some concave profit function for the firm, the equivalent of the standard risk aversion that justifies insurance in the general case. Thus, for instance, we could model as well the shocks through random values of $G$. 
$R_{I_{2}}^{H_{1}} = G$. When the bank is not committed to lend the contract is characterized by $R_{I_{2}}^{H_{1}} = \max(G, \frac{1+r_{H}}{p_{G}})$.

**Proof.** Assume, by way of contradiction a contract $(R_{I_{1}}^{I_{1}}, R_{I_{2}}^{I_{1}}, R_{I_{2}}^{H_{1}})$ that is such that $R_{I_{2}}^{H_{1}} > G$. Then for some $\varepsilon > 0$, the contract $(R_{I_{1}}^{I_{1}} + \varepsilon, R_{I_{2}}^{I_{1}}, R_{I_{2}}^{H_{1}} - \frac{\varepsilon}{(1-p_{G})})$ will be preferred by the borrowing firms. This is the case as the difference in the cost of funding will be

$$\Delta C = \varepsilon - (1-p)_{G} \left[ C \left( R_{I_{2}}^{H_{1}} - G \right) - C \left( R_{I_{2}}^{H_{1}} - \frac{\varepsilon}{(1-p_{G})} - G \right) \right] = \varepsilon (1-t)$$

for some $Z = R_{I_{2}}^{H_{1}} - \frac{\varepsilon}{(1-p_{G})}, \theta (0,1)$.

Because of $C'(0) > 1$ jointly with the convexity of $C$, we know $C'(Z) > 1$ and therefore, $\Delta C < 0$ proving the first part of the proposition.

Regarding the second part, if the bank is allowed not to lend, it will lend

only if this provides a higher return than the market riskless rate, so that

the bank has incentives to lend if and only if the constraint $p_{G}R_{I_{2}}^{H_{1}} \geq r_{H}$ is satisfied.

Notice that when $R_{I_{2}}^{H_{1}} = \frac{1+r_{H}}{p_{G}}$ there is still some degree of insurance, as the bank is giving up its market power as well as its ex post monopoly profits when interest rates are high in order to have higher profits when interest rates are low. This point is emphasized in the following corollary.

**Corollary 5** When the bank is not committed to lend, the amount of insurance increases with the monitoring cost and with market power. In the extreme case of competition with zero monitoring cost, the banks cannot provide insurance against interest rate shocks.

When the bank cannot commit to lending, the amount of insurance the bank provides will be the difference between the second period rate, $(1+\rho)(1+r_{H}+M)$ and $\max(G, \frac{1+r_{H}}{p_{G}})$. Therefore, the amount of insurance equals $\min(\frac{(1+\rho)(1+r_{H}+M)}{p_{G}} - G, \frac{m+1+r_{H}(\rho)}{p_{G}})$, so it depends positively on $\rho$ and on $M$. In the extreme case when the market is competitive and monitoring costs are zero, there is no room for intertemporal insurance.

**Remark 1** When the firm is able to renegotiate (lenient bank case), the bank is unable to provide any intertemporal smoothing.
This is indeed the case, as the firm would renegotiate for low interest rates and not for high ones, creating a situation where the bank is always at a loss.

Berlin and Mester (2000) combine the banks interest rate insurance services with the fact that they have to break even at each period. By so doing they predict that banks with more core deposits (whose remuneration is close to zero) are able to provide more insurance than those which are funded with liabilities yielding the market interest rate. The empirical evidence they obtain backs up their model’s conclusions.

Even if there is no clear theoretical justification for imposing a period-by-period no-loss constraint on the banks, it could be argued that the volatility of profits is costly for the banks as it increases their cost of capital. This is sufficient for Berlin and Mester prediction to hold and those banks with more access to core deposits will provide more intertemporal smoothing.

Notice that the perspective could be reversed, if we consider the supply for core deposits. In this case it will be the banks providing more intertemporal smoothing that will be willing to pay more (in non-pecuniary services) to attract core deposits.

6. Conclusion

In the present paper we revisit the main well established results on relationship banking using a different model. In this way, we hope to achieve two aims at the same time: to survey the main results in a synthetic way and to test their robustness. We have focussed on three main types of results related to relationship banking: intertemporal pricing, the effect of competition on relationship banking and intertemporal insurance.

One of the main predictions of the relationship banking model is that banks will initially set below-cost prices (rates) and make up with profits generated by their ex post monopoly rents at latter stages. Our perspective on this classic and robust result brings in a caveat, that the firm should not be able to make a take it or leave it offer to the bank, as this would eliminate the ex post monopoly situation.

Our examination of the effect of competition on relationship banking shows mainly that there are two different issues at stake. The first one is the one related to the investment the bank makes in the rela-
tionship, and it is higher the higher the market power the bank has. The second one is related to the total amount of relationship loans, as opposed to direct finance, and this depends upon how competitive the two markets are. The implication here is that modeling competition in relationship banking may involve the use of a more precise definition of “competition”. The consequences of increased competition are completely different if we refer to ex-ante or to ex-post competition, as in the second case the ex post monopoly rents disappears. So this is clearly a direction for future research.

Regarding the effects of relationship banking on intertemporal insurance, our model shows that this is quite a robust result. It is indeed the case that in equilibrium banks and firms are better off under this type of scheme. Still, there is a limit to the amount of intertemporal smoothing a bank can produce, as it may have alternative investment options. If the bank is not committed to lend, it is easy to show that the larger the ex post monopoly the larger the level of intertemporal smoothing.

References


Resumen

Durante la última década, el concepto de banca de relación ha cobrado protagonismo en el área de la teoría bancaria, reflejando la importancia de compartir información y los beneficios de una relación continuada. Las contribuciones teóricas han examinado las implicaciones de la banca de relación para la estructura del mercado bancario, pero sus conclusiones son a veces contradictorias. El presente artículo constituye una panorámica analítica donde, gracias a un modelo básico sencillo, se estudia las implicaciones de la banca de relación para el precio de los préstamos así como para el grado de competencia en el sector bancario.

Palabras clave: Banca de relación, competencia, estructura de mercado.

Recepción del original, febrero de 2002
Versión final, julio de 2004