The Structure of Bank Relationships, Endogenous Monitoring and Loan Rates

Elena Carletti*
University of Mannheim, Department of Economics,
D-68131 Mannheim, Germany
Email: carletti@rumms.uni-mannheim.de

March 30, 2003

Abstract

This paper investigates a firm's choice between borrowing from a single bank and from two banks. The focus is on how this decision affects banks' equilibrium monitoring intensities and loan rates. Two-bank lending suffers from duplication of effort and sharing of monitoring benefits, but it benefits from diseconomies of scale in monitoring. Thus, two-bank lending involves lower monitoring but not necessarily higher loan rates than single-bank lending. The optimal borrowing structure balances the benefit of monitoring for the firm in terms of higher success probability of the project against its drawbacks of lower expected private return and higher total monitoring costs. In contrast to the previous theoretical literature, the model lays down an explanation for the empirical observation that multiple-bank lending does not unambiguously increase loan rates or firms' quality, in particular in small business lending.

JEL Classification: C72, D82, G21, G32

Keywords: single-bank lending, multiple-bank lending, monitoring, loan rates, small business lending

*I am particularly grateful to Margaret Bray for her encouragement, to Martin Hellwig for his suggestions, and to an anonymous referee for constructive criticism that helped to improve the paper. I also appreciated comments and suggestions from Franklin Allen, Angelo Baglioni, Vittoria Cerasi, Marco Da Rin, Mariassunta Giannetti, Antoine Faure-Grimaud, Xavier Freixas, Antonio Guarino, Philipp Hartmann, Heidrun Hoppe, Andreas Irmen, Christian Laux, Robert Marquez, Annamaria Menichini, Marco Polenghi, Jorn Rothe, Dezsoe Szalay, Giancarlo Spagnolo, the managing editor Anjan Thakor, Xavier Vives, seminar participants at Ente Einaudi, LSE, Mannheim, Munich and Bologna University, ESEM 2001, as well as the Fed New York/JFI/Stern Businesss School Symposium on 'Corporate Governance in Banking and Financial Services Industries'. All errors are, of course, my own.

1 Introduction

This paper addresses the question of why in many countries even relatively small firms borrow from more than one bank.[1] Modern theories of financial intermediation relate the benefits of multiple-bank lending to the inefficiencies affecting exclusive bank-firm relationships, namely the hold-up and the soft-budget-constraint problems. As for the hold-up problem, Sharpe (1990) and Rajan (1992) argue that a relationship bank may use the superior private information it possesses about the firm to extract rents, thus distorting entrepreneurial incentives and causing inefficient investment choices. In this context, borrowing from multiple banks can restore competition among banks and, consequently, improve entrepreneurial incentives (Von Thadden, 1992; Padilla and Pagano, 1997). As for the soft-budget-constraint problem, Dewatripont and Maskin (1995) argue that a relationship bank may refinance unprofitable projects and thus reduce entrepreneurial incentives to prevent default. By complicating the refinancing process and making it less profitable, multiple-bank lending allows banks to commit not to extend further inefficient credit. Similarly, Bolton and Scharfstein (1996) show that multiple-bank lending reduces entrepreneurial incentives to default strategically because it complicates debt renegotiation.[2]

The theories mentioned above predict that firms borrowing from multiple banks should represent better risks and should pay lower loan rates than firms borrowing from a single bank. These predictions, however, are not entirely consistent with the empirical evidence. Petersen and Rajan (1994) find that in the US multiple-bank lending increases loan rates substantially but has no role in explaining firms' quality. Studies on Germany (Harhoff and Körting, 1998; Elsas and Krahnen, 1998) show that multiple-bank lending does not affect loan rates but is negatively correlated with firms' financial status. Studies on Italy (Foglia *et al.*, 1998; D'Auria *et al.*, 1999) find that borrowing from multiple banks reduces both loan rates and firms' quality.[3]

How does one explain the discrepancy between the theories based on the hold-up and the soft-budget-constraint problems and the empirical evidence? One explanation, as I see it, is that the theoretical literature has not explicitly considered banks' incentives to monitor.

[1] Cross-country evidence on the number of bank relationships is provided in Ongena and Smith (2000a) and Ongena and Smith (2000b).

[2] Other rationales for multiple-bank lending refer to firms' desire to reduce liquidity risk and disclose information through credit relationships. Detragiache *et al.* (2000) argue that, when relationship banks face internal liquidity problems, borrowing from multiple banks can avoid early liquidation of profitable projects. Yosha (1995) shows that firms may prefer multiple-bank lending as a way to disclose confidential information about the quality of their projects and to avoid aggressive behavior by competitors. See also Boot (2000) for a recent survey on relationship banking.

[3] Degryse and Ongena (2001) find evidence of a negative correlation between the number of bank relationships and firms' quality in Norway.

Most contributions assume that banks acquire information about their borrowers as a simple by-product of the lending activity (e.g., Padilla and Pagano, 1997), or that monitoring is costly but its level is exogenous. In the latter case, a bank monitors with the same intensity irrespective of whether it is the sole lender or finances the firm with other banks (e.g., Von Thadden, 1992).[4] This paper contributes to the literature on multiple-bank lending by analyzing how the number of bank relationships influences banks' monitoring incentives and how, in turn, the level of monitoring affects loan rates and a firm's choice between borrowing from a single bank versus borrowing from two banks.

I address these issues in a simple one-period model, which builds on Holmström and Tirole (1997). An entrepreneur in need of external funds has to decide whether to exert effort and to increase the success probability of a risky project. Not exerting effort earns the entrepreneur a private benefit, but makes credit unavailable. The firm's moral hazard problem can be ameliorated through bank monitoring, which induces the entrepreneur to exert effort. Monitoring is, however, costly and not contractable. As a consequence, a bank chooses the level of monitoring so as to maximize its profit rather than the total surplus of the project. A bank's incentive to monitor depends on whether it is the sole lender or it finances the firm with another bank. Two-bank lending suffers from duplication of effort and sharing of monitoring benefits, but it benefits from diseconomies of scale in monitoring. Thus, two-bank lending involves lower monitoring but not necessarily higher loan rates than single-bank lending.

In choosing the number of banks it should borrow from, the firm balances the different effects that monitoring has on its profit. On the one hand monitoring benefits the firm by increasing the success probability of the project, but on the other hand it reduces the firm's expected private return and increases total monitoring costs. Since banks do not internalize the firm's private benefit, they may choose a monitoring intensity which, from the firm's perspective, is excessive. When overmonitoring is an issue, the firm may prefer to borrow from two banks as a way to reduce overall monitoring, although this may imply higher loan rates. When overmonitoring is not an issue, the firm may prefer single-bank lending, since it avoids duplication of effort and free-riding problems, and lowers monitoring costs. Which borrowing structure is optimal will depend on the size of the cost of monitoring and of the private benefit, as well as on the firm's expected profitability. The model predicts that the attractiveness of two-bank lending is increasing in the cost of monitoring, the private benefit and the firm's expected profitability. These predictions find empirical support in Ongena and Smith (2000b) and Detragiache et al. (2000).

[4] An exception is Dewatripont and Maskin (1995), who endogenize the level of bank monitoring, but analyze only the case of a single monitor.

This basic framework is extended in three directions. First, I consider a linear monitoring cost structure. In this case, two-bank lending always implies higher loan rates than single-bank lending, but it may still be optimal as a way to reduce overall monitoring. Second, I consider the possibility of the firm borrowing from a single bank and other outside investors. By choosing the amount to borrow from the bank, the firm can reduce the bank's incentive to monitor and induce levels of monitoring that are not attainable via a direct contract with the bank. Thus, borrowing from a single bank and other outside investors may lead to less monitoring than borrowing from two banks, and it may be optimal when the firm is highly profitable and misbehavior is accompanied by a high private benefit. Third, I analyze an alternative model of bank control based on Rajan and Winton (1995) and Park (2000), where monitoring leads to a threat of liquidation in case of misbehavior. Such a threat improves the firm's incentive to behave well, but it does not change the firm's optimal borrowing choice.

The novelty of this work is that it analyzes how banks' incentives to monitor change with the number of banks the firm borrows from, and how this affects the firm's optimal borrowing choice. The results are particularly applicable to the financing of small businesses, since the model is set up so that a bank does not lend to the firm if it does not monitor and there is no public information about the firm's quality. In this respect, the paper provides a new explanation for multiple-bank lending, which contrasts sharply with the hold-up literature.

This paper is related to a number of others. Thakor (1996) analyzes the optimal number of banks a firm approaches for credit in a model where banks perform screening to sort out borrowers' creditworthiness. Approaching more banks is beneficial in that it increases the probability of receiving credit and lowers the loan rate. It is costly, however, in that it reduces each bank's incentive to screen, thus increasing the probability of being denied credit. This trade-off limits the attractiveness of multiple banks, but approaching more than a single bank is always optimal. By contrast, I focus on the number of banks a firm should borrow from in a model where banks perform postlending monitoring and their incentives depend on the number of banks that finance the firm. In this context, multiple-bank lending is not always optimal, and its attractiveness depends on the level of monitoring as well as on the cost of financing.

Other papers analyze banks' incentives to monitor. Besanko and Kanatas (1993) rely on the non-contractability of bank monitoring to explain the coexistence of banks and capital markets. Covitz and Heitfield (1999) focus on banks' monitoring incentives to analyze the relationship between market power, loan rates and bank risk. Cerasi and Daltung (2000) show that diversification can improve banks' incentives to monitor, when they are debt financed. All these models assume that firms borrow from a single bank, which is also the sole monitor. By contrast, I look at how banks' incentives to monitor change with the number of banks the firm borrows from.

In this respect, the paper is also related to some contributions in the literature on shareholders' monitoring. Winton (1993) analyzes the monitoring decisions of multiple shareholders, and suggests that sufficiently convex costs can make multiple monitors more efficient. He does not address, however, the firm's choice of the optimal number of monitoring shareholders. Pagano and Röell (1998) suggest that, when a firm has the possibility to enjoy high private benefits, it may prefer to go public and limit the stake of the monitoring shareholder in order to reduce outside monitoring. In a similar spirit, I consider the possibility of the firm borrowing from a single bank and other outside investors as a way to reduce bank monitoring. However, while Pagano and Röell (1998) examine the trade-off between the cost of listing and overmonitoring, I focus on the firm's choice between borrowing from a single bank and outside investors and from two banks. The idea of excessive outside monitoring is also seen in Burkart *et al.* (1997) in the context of optimal ownership dispersion. Other related contributions are Rajan and Winton (1995) and Park (2000), who concentrate on how to structure debt contracts so as to maximize lenders' incentives to monitor.

The remainder of the paper is organized in four sections. Section 2 describes the basic model. Section 3 derives the equilibrium with bank monitoring. Section 4 discusses the robustness of the basic model. Section 5 contains the concluding remarks and the empirical predictions.

2 The Basic Model

Consider a two-date economy ($T = 0, 1$) with two classes of risk neutral agents: a single firm and a perfectly competitive banking sector. The firm has access to a risky project, and needs external funds to finance it. Only bank lending is available, and the firm can borrow from a single bank or from two banks.[5]

The project requires an investment of one unit at date 0, and yields a pecuniary return of R if it succeeds and 0 if it fails at date 1. The success probability of the project depends on the behavior of the entrepreneur running the firm: it is p_H if he behaves well ('high' effort) and p_L if he misbehaves ('low' effort), with $p_H > p_L$. The project is creditworthy only in case of good behavior, i.e., $p_H R > 1 > p_L R$. The entrepreneur may choose to misbehave, however, in order to enjoy a non-transferable private benefit B. This can be broadly interpreted as a quiet life, diversion of corporate revenues for private use, unprofitable pet projects, advantageous deals with other firms and also managerial perks from project completion. There is a moral hazard problem because the entrepreneur's behavior is not observable.

[5] Other sources of financing are considered in Section 4.

Banks raise funds at the riskless interest rate (assumed to be zero, for simplicity), and lend to the firm only if they expect non-negative profits, i.e., only if the firm behaves well. This occurs if and only if

$$p_H(R-r) \geq p_L(R-r)+B,$$

where r is the (gross) per unit loan rate. To provide a role for bank monitoring, I assume that simple lending is not feasible; i.e.,

$$\frac{(p_H - p_L)}{p_H}(p_H R - 1) < B. \qquad (1)$$

Assumption (1) means that the private benefit is sufficiently high to induce the firm to misbehave, when the loan rate is set at the competitive level $\frac{1}{p_H}$. Thus, Assumption (1) implies that the project is not financed if banks simply lend to the firm.[6]

Suppose now that banks can mitigate the firm's moral hazard problem through monitoring. In particular, monitoring allows banks to observe the firm's behavior, and to intervene if it misbehaves.[7] Each bank chooses to monitor with an intensity $M \in [0,1]$, which determines the probability with which each bank is able to improve the firm's behavior. Monitoring is costly; an intensity M costs $C(M) = \frac{mM^2}{2}$. The convex cost function reflects the idea that it is increasingly difficult for a bank to find out more and more about a firm. Also, a bank has scarce human resources for adequate monitoring, and cannot expand them indefinitely.[8] As a consequence, monitoring exhibits diseconomies of scale, whose importance is measured by the parameter m (henceforth, also referred to as "cost of monitoring").[9]

A bank's monitoring intensity M is not observable and thus not contractable. This introduces another moral hazard problem in the model, and it may lead to either an overmonitoring or an undermonitoring problem with respect to the first-best monitoring level M^{FB}. In fact, whereas M maximizes a bank's expected profit, M^{FB} maximizes the total return of the project−defined as the financial *and* the private return of the project−and is

[6] In essence, Assumption (1) creates a role for bank monitoring. If it was not satisfied, banks would act only as lenders, and the firm would be indifferent between borrowing from a single bank and two banks.

[7] The idea is similar to the one in Holmström and Tirole (1997). Monitoring brings down the private benefit that the firm enjoys from misbehaving, and therefore it induces it to behave well. This monitoring technology simplifies the analysis—in particular, the firm's behavioral choice—without affecting the qualitative results of the model, as discussed in Section 4.

[8] For example, banks are often unable to hire additional skilled loan officers. Alternatively, diverting internal staff to the monitoring of a specific project may require additional training or result in an indirect cost through the negligence of other duties. On this latter explanation, see also Cerasi and Daltung (2000).

[9] Although assuming a convex monitoring cost function appears rather plausible, a linear monitoring cost function is considered in Section 4. Moreover, note that the convex cost function ignores the possibility of some initial fixed costs in learning about a firm, since $C(0) = C'(0) = 0$. Extending the model to the case $C'(0) > 0$ is straightforward. The results would be somewhere between those obtained with the convex cost function and those obtained with the linear cost function in Section 4.

equal to:

$$M^{FB} = \begin{cases} 1 & \text{if } m \leq \overline{m}_{FB} = (p_H - p_L)R - B \\ \frac{(p_H - p_L)R - B}{m} & \text{if } m > \overline{m}_{FB}. \end{cases}$$

The intensity M^{FB} depends on the cost of monitoring; it is equal to one if m is sufficiently low ($m \leq \overline{m}_{FB}$) and is less than one otherwise ($m > \overline{m}_{FB}$). To ensure that a positive level of monitoring is always optimal, I assume that M^{FB} is positive, i.e.,

$$B < (p_H - p_L)R. \tag{2}$$

Assumption (2) puts an upper bound on the level of private benefit B, and it implies that the firm would behave well if it did not need to raise external funds.

The timing of the model is as follows. At the beginning of date 0 the firm chooses whether to borrow from a single bank or from two banks; its choice is observable. Then the firm contacts banks, and a two-stage game starts. In the single-bank case, the firm approaches a single bank, and offers it a contract specifying the loan rate. The bank can either accept or reject the contract. If the bank rejects it, the game ends. If the bank accepts the contract, the project is financed and then simultaneously the firm chooses its behavior and the bank chooses its monitoring intensity.[10] At date 1 the project matures and claims are settled. If the project succeeds, the firm pays the loan rate to the bank and retains the surplus. The two-bank case has the same time structure. Figure 1 summarizes the timing of the model if the contract is accepted.

T=0			T=1
the firm chooses whether to borrow from a single bank or two banks	the firm offers loan contract to the bank(s)	the firm chooses p_H or p_L; bank(s) chooses M	project matures; claims are settled

Figure 1: Timing of the model.

3 The Competitive Equilibrium with Bank Monitoring

The model is solved as follows. First, I take the firm's choice of the number of banks as given, and I characterize the equilibrium of the game with a single bank and with two banks.

[10] The choice of simultaneous moves between the firm and the bank does not affect the results as long as agents' decisions are not observable. The results are affected only if the bank chooses its monitoring intensity first, and the firm can observe it before moving. This order of moves would correspond to the case in which the bank can commit to a specific monitoring intensity.

Second, I analyze the firm's borrowing choice, taking the equilibrium of the game in each scenario as given.

3.1 Single-Bank Financing

In this section I characterize the equilibrium of the game with a single bank (henceforth, "single-bank game"). Because the banking system is perfectly competitive, the bank accepts the contract only if it expects to break even and the contract is feasible. This means that the loan rate must satisfy the bank's zero-profit condition and it must not exceed the pecuniary return of a successful project. After lending has occurred, the firm and the bank choose their actions so as to maximize their expected profits. Formally, a competitive equilibrium of the single-bank game in which the project is financed is a subgame-perfect equilibrium such that:

1) the firm sets the loan rate so as to maximize its expected profit subject to the bank's zero-profit condition and the contract feasibility condition, correctly anticipating its behavioral choice and the bank's monitoring intensity;

2) the firm's behavioral choice and the bank's monitoring intensity constitute a Nash equilibrium of the behavior/monitoring subgame.

Let r_1 be the loan rate and M_1 be the bank's monitoring intensity. The firm's expected profit is:

$$\Pi_{F_1}^H = p_H(R - r_1) \tag{3}$$

when it behaves well, and

$$\Pi_{F_1}^L = M_1 p_H(R - r_1) + (1 - M_1)[p_L(R - r_1) + B] \tag{4}$$

when it does not. The bank's expected profits are instead given by:

$$\Pi_{B_1}^H = p_H r_1 - 1 - \frac{m}{2}(M_1)^2, \tag{5}$$

$$\Pi_{B_1}^L = M_1 p_H r_1 + (1 - M_1)p_L r_1 - 1 - \frac{m}{2}(M_1)^2 \tag{6}$$

when the firm behaves well and when it does not.

Proposition 1 characterizes the competitive equilibrium of the single-bank game.[11] Definition 1 introduces the function that will describe the equilibrium loan rate.

Definition 1 Let $g_1(m)$ be the following function:

$$g_1(m) = \begin{cases} \frac{2+m}{2p_H} & \text{if } m \leq \overline{m}_1 = \frac{2(p_H - p_L)}{p_H + p_L} \\ \frac{-mp_L + \sqrt{m^2 p_L^2 + 2m(p_H - p_L)^2}}{(p_H - p_L)^2} & \text{if } m > \overline{m}_1. \end{cases}$$

[11] All proofs are in the appendix.

Proposition 1 *There exists a unique competitive equilibrium of the single-bank game in which the project is financed if and only if its return, R, satisfies the condition $R \geq g_1(m)$. The equilibrium has the following features:*
1) the loan rate is $r_1^* = g_1(m)$;
2) the firm misbehaves; there is a critical level of the cost of monitoring, $\overline{m}_1$, such that the bank monitors with intensity $M_1^* = 1$ if $m \leq \overline{m}_1$, and with $M_1^* < 1$ otherwise. More precisely,

$$M_1^* = \begin{cases} 1 & \text{if } m \leq \overline{m}_1 \\ \frac{-mp_L + \sqrt{m^2 p_L^2 + 2m(p_H - p_L)^2}}{m(p_H - p_L)} & \text{if } m > \overline{m}_1. \end{cases} \tag{7}$$

Proposition 1 shows the role of bank monitoring in the model. As already mentioned, under Assumption (1) the project is not financed in the absence of monitoring, since the bank would not break even. The project can instead be financed in the presence of monitoring, since this induces the firm to behave well, thereby increasing the success probability of the project.[12]

The bank always exerts a positive intensity of monitoring, since the convexity of the monitoring costs implies $C'(0) = 0$. The equilibrium intensity equals the first-best level M^{FB} if $m \leq \overline{m}_{FB}$ and is higher otherwise.[13] Thus, single-bank lending implies an overmonitoring problem if $m > \overline{m}_{FB}$. The reason is that the bank chooses the monitoring intensity to maximize its own expected profit, which does not take the firm's private benefit into account.

The contract feasibility condition, $r_1^* \leq R$, requires that the project is sufficiently profitable and that monitoring is not too costly. This implies the following corollary.

Corollary 1 *In the competitive equilibrium, the project is financed if and only if its return and the cost of monitoring satisfy the conditions $R \geq \frac{2}{p_H + p_L}$ and $m \leq \frac{(p_H - p_L)^2 R^2}{2(1 - p_L R)}$.*

3.2 Two-Bank Financing

In this section I characterize the equilibrium of a game with two (identical) banks (henceforth, "two-bank game"), and I compare it with the equilibrium of the single-bank game. As before, a competitive equilibrium of the two-bank game in which the project is financed is a subgame-perfect equilibrium such that the loan rate satisfies each bank's zero profit condition and the

[12] Assumption (1) also implies that the firm still chooses to misbehave when it is not monitored. Thus, the firm's incentive does not improve in the presence of monitoring. Rather, it worsens through the higher loan rate needed to cover the monitoring costs. This particular effect has already been encountered in different contexts (e.g., Rajan, 1992; Padilla and Pagano, 1997).

[13] Assumption (1) implies $M_1^* > M^{FB}$ if $m > \overline{m}_{FB}$, even if $r_1^* < R$.

feasibility constraint; after lending has occurred, the firm and the two banks choose their actions so as to maximize their expected profits.

The contrast between the single-bank game and the two-bank game depends crucially on how the two banks interact in their monitoring decisions. I assume that they lend half a unit each to the firm, and choose how much to monitor simultaneously and independently. Each bank's monitoring intensity, however, has an overall impact on the firm's behavior. When one bank discovers that the firm misbehaves, it induces the firm to behave well on the *whole* project. The idea is that the decision to undertake monitoring is private information, but its outcome is publicly known. Let r_2 be the loan rate in the two-bank game, M_i be bank i's monitoring intensity with $i = A, B$, and $\overline{M}_2$ be the *overall* monitoring intensity, which is equal to

$$\overline{M}_2 = M_A + M_B - M_A M_B. \tag{8}$$

The firm's expected profits when it behaves well and when it does not are given by:

$$\Pi^H_{F_2} = p_H(R - r_2), \tag{9}$$

$$\Pi^L_{F_2} = \overline{M}_2 p_H(R - r_2) + (1 - \overline{M}_2)[p_L(R - r_2) + B]. \tag{10}$$

The expected profit of each bank i is:

$$\Pi^H_{B_i} = p_H\frac{r_2}{2} - \frac{1}{2} - \frac{m}{2}(M_i)^2 \tag{11}$$

when the firm behaves well, and

$$\Pi^L_{B_i} = \overline{M}_2 p_H\frac{r_2}{2} + (1 - \overline{M}_2)p_L\frac{r_2}{2} - \frac{1}{2} - \frac{m}{2}(M_i)^2 \tag{12}$$

when it does not. Expressions (8), (11) and (12) highlight the features of the two-bank game. First, the two banks face an externality and a duplication of effort (second and third term in (8)), since each bank's monitoring affects the whole project and is not observable. Second, they share the monitoring benefits because each of them gets $\frac{r_2}{2}$, but pays the full cost of monitoring. Third, the two banks benefit from the diseconomies of scale due to the convexity of the cost function.

Proposition 2 characterizes the competitive equilibrium of the two-bank game. Definition 2 introduces the function that will describe the equilibrium loan rate.

Definition 2 Let $g_2 = g_2(m)$ be the unique positive real solution of the following equation: $f(g_2) = p_H(p_H - p_L)^2 g_2^3 + (p_H - p_L)(3mp_H + mp_L - (p_H - p_L))g_2^2 - 4m((p_H - p_L) - mp_L)g_2 - 4m^2 = 0$.

Proposition 2 *There exists a unique symmetric competitive equilibrium of the two-bank game in which the project is financed if and only if its return, R, satisfies the condition $R \geq g_2(m)$. The equilibrium has the following features:*
1) the loan rate is $r_2^* = g_2(m)$;
2) the firm misbehaves; each bank monitors with intensity

$$M_2^* = \frac{(p_H - p_L)r_2^*}{(p_H - p_L)r_2^* + 2m} < 1. \tag{13}$$

Similar to the single-bank game, Proposition 2 shows that the project can be financed in the presence of monitoring. The two banks monitor the firm with the same positive intensity M_2^*.[14] The denominator in expression (13) reflects two of the earlier-mentioned features of the two-bank game. The term $2m$ comes from the sharing of monitoring benefits; the term $(p_H - p_L)r_2^*$ reflects the duplication of effort. Both of these elements curtail the two banks' incentives to monitor; consequently, they always monitor with an intensity lower than one.

The equilibrium loan rate r_2^* depends on how the duplication of effort and the sharing of monitoring benefits interact with the convexity of the cost function. This interaction works through the level of M_2^*. Given the form of the function $f(g_2)$, the analytical expression for r_2^* is quite difficult to handle. This complicates also the comparison between the equilibrium loan rates of the single-bank and the two-bank game. Formally, from (6) and (12), r_1^* and r_2^* are given by:

$$r_1^* = \frac{1 + C(M_1^*(r_1^*))}{[p_L + M_1^*(r_1^*)(p_H - p_L)]}, \tag{14}$$

$$r_2^* = \frac{1 + 2C(M_2^*(r_2^*))}{[p_L + \overline{M}_2^*(r_2^*)(p_H - p_L)]}. \tag{15}$$

Thus, the equilibrium loan rates are implicitly determined by the ratio between total monitoring costs and the total success probability of the project. By using the Implicit Function Theorem, one can show that both r_1^* and r_2^* increase with the cost of monitoring m, and both decrease with the success probabilities p_H and p_L.[15] Also, one can show that r_2^* is higher than r_1^* in the neighborhood of $m = 0$. Whether this still holds as m increases depends on how total monitoring costs and the total success probability of the project in (14) and (15) change with the parameters of the model. Intuitively, it depends on whether the two

[14] In Proposition 2 I focus on symmetric equilibria. Note, however, that, if $(p_H - p_L)r_2 > 2m$, there exists also an asymmetric equilibrium in which only one bank monitors, i.e., $M_i = 1$ and $M_j = 0$ for $i \neq j$. This equilibrium is 'equivalent' to the one of the single-bank game with $M_1^* = 1$. The two equilibria always coexist, and lead to the same monitoring intensities and loan rates.

[15] Proofs are available on request.

banks' advantage arising from the diseconomies of scale dominates the drawback of lower monitoring incentives stemming from the duplication of effort and the sharing of monitoring benefits. To see this, I proceed as follows. I start with comparing monitoring intensities and total monitoring costs in the single-bank and the two-bank game under the assumption that the loan rate is the same in the two cases. Next, I use numerical simulations to characterize the equilibrium loan rates, and I show that the single bank does not always require a lower loan rate than the two banks.

Suppose for a moment that the two banks charge the same loan rate as the single bank, i.e., $r_2^* = r_1^* = r$. Figures 2 and 3 illustrate how monitoring intensities and total monitoring costs in the single-bank and the two-bank game change as one increases the cost of monitoring m in the interval $[0, 1]$. Lemmas 1 and 2 follow immediately.

Figures 2 and 3 about here

Lemma 1 *Let the loan rate be the same in the single-bank and the two-bank game, i.e., $r_1 = r_2 = r$. Then the monitoring intensity of the single bank is always higher than the overall monitoring intensity of the two banks.*

Lemma 2 *Let the loan rate be the same in the single-bank and the two-bank game, i.e., $r_1 = r_2 = r$. Then the monitoring costs of the single bank are lower than the total monitoring costs of the two banks if the cost of monitoring is $m < \frac{(p_H - p_L)r}{2(1+\sqrt{2})}$, and higher otherwise.*

The main intuition behind Lemma 1 is that the duplication of effort and the sharing of monitoring benefits curtail the two banks' monitoring incentives so much that they monitor less than the single bank even in the *aggregate*. The difference between the monitoring intensities of the two games, however, is not monotonic in the cost of monitoring. As Figure 2 shows, it decreases if $m \leq (p_H - p_L)r$, and increases otherwise. The reason is that, as m increases beyond $(p_H - p_L)r$, the two banks reduce monitoring less than the single bank because the duplication of effort becomes smaller. Thus, Lemma 1 suggests that a firm is of a better quality–in terms of a higher total success probability of the project–if it borrows from a single bank. The relative advantage of borrowing from a single bank is more pronounced for intermediate levels of the cost of monitoring, and it vanishes towards very low and very high levels.

Lemma 2 suggests that the difference between the monitoring costs of the single-bank and the two-bank game depends on whether the duplication of effort overcomes the diseconomies of scale. If the cost of monitoring is very low, each of the two banks monitors almost as much as the single bank, but in the aggregate they face twice the costs because of the large duplication of effort and the small diseconomies of scale. As m increases, the two banks reduce monitoring and benefit from the larger diseconomies of scale. As m reaches $\frac{(p_H - p_L)r}{2(1+\sqrt{2})}$, these two effects become so important that the two banks have lower total monitoring costs than the single bank. Their relative advantage starts to decrease as m increases beyond $(p_H - p_L)r$ because from that point onwards the two banks reduce monitoring less than the single bank.

The relationship between the equilibrium loan rates r_1^* and r_2^* hinges on the relative difference between expressions (14) and (15). Lemmas 1 and 2 suggest that both total monitoring costs and the total success probability of the project are higher in (14) than in (15) if $m > (p_H - p_L)r$, and that their relative differences decline as m increases beyond this level. The final effect is somewhat ambiguous, and it depends on the success probabilities p_H and p_L. To see this, I use numerical simulations. I fix $p_H = 1$, and examine the behavior of the loan rates as m increases in the interval $[0, 1]$, first for $p_L = 0.4$ and then for $p_L = 0.6$. Figures 4 and 5 illustrate the results of the simulations.

Figures 4 and 5 about here

The figures show that the two banks always require a higher loan rate than the single bank when $p_L = 0.4$, but not when $p_L = 0.6$. As Figure 5 illustrates, when $p_L = 0.6$ there is a critical level of the cost of monitoring, $\widehat{m} \in [\overline{m}_1, 1]$, such that $r_2^* < r_1^*$ if $m > \widehat{m}$. Formally, the reason is that, as both m and p_L increase sufficiently, the smaller difference between the total success probabilities in (14) and (15) dominates the one in total monitoring costs. Intuitively, as both m and p_L increase sufficiently, the two banks become more efficient than the single bank because the diseconomies of scale dominate the duplication of effort and the sharing of monitoring benefits. This result holds only when $p_L = 0.6$. Indeed, increasing p_L reinforces the effects of increasing m; it reduces monitoring more in the two-bank case than in the single-bank case, and it has a greater direct positive impact on the total success probability of the project in (15) than in (14) because $\overline{M}_2^* < M_1^*$.

The parameter p_L—the success probability of the project in case of misbehavior—can be interpreted as a measure of the severity of the firm's moral hazard problem. Next, I disentangle this problem into a *financial* and a *private* problem. The former, measured by p_L, concerns the success probability of the project if the firm misbehaves; the latter, captured by the private benefit B, relates to the private return that the firm enjoys from misbehaving. The two problems have different effects in the model. The financial moral hazard problem affects banks' monitoring incentives and loan rates, whereas the private moral hazard problem influences only the firm's choice of the number of banks, as shown in the next section. With this distinction in mind, Proposition 3 summarizes the behavior of the equilibrium loan rates of the single-bank and the two-bank game. The result provides a theoretical rationale for the empirical observation that borrowing from a single bank does not unambiguously lead to a lower cost of financing.

Proposition 3 *When the firm's financial moral hazard is weak, the two banks require a lower loan rate than the single bank if the cost of monitoring is sufficiently high. In all other circumstances, the two banks require a higher loan rate than the single bank.*

3.3 The Firm's Choice between a Single Bank and Two Banks

I turn now to the firm's choice between borrowing from a single bank (henceforth, "single-bank" lending) and borrowing from two banks (henceforth, "two-bank" lending). Once the respective equilibrium monitoring intensities and loan rates are substituted in (4) and (10), the firm's expected profits are given by:

$$\Pi^L_{F_1} = [M_1^* p_H + (1 - M_1^*) p_L] R + (1 - M_1^*) B - [1 + \frac{m}{2} (M_1^*)^2], \tag{16}$$

$$\Pi^L_{F_2} = [\overline{M}_2^* p_H + (1 - \overline{M}_2^*) p_L] R + (1 - \overline{M}_2^*) B - [1 + m (M_2^*)^2] \tag{17}$$

if the firm borrows from a single bank and two banks, respectively. The terms on the right hand side of (16) and (17) represent, in order, the expected financial return of the project, the firm's expected private return and the banks' expected repayment, which is equal to the sum of the amount of loan and total monitoring costs.

The firm chooses the borrowing structure that maximizes its expected profit. Its choice depends on the relative differences between monitoring intensities and loan rates –or equivalently total monitoring costs– in the single-bank and the two-bank game. Thus, the firm must balance the benefit of monitoring in terms of higher expected financial return of the project against the drawbacks of lower expected private return and higher total monitoring costs. Once again, in order to analyze the firm's optimal choice, I use numerical simulations. I fix $R = 1.6$, $p_H = 1$, $p_L = 0.4$, and examine the firm's profits in (16) and (17) as m increases in the interval $[0, 1]$ and B increases from 0.4 to 0.6.[16] Figure 6 graphs the results of the simulations.

Figure 6 about here

Figure 6 shows that the firm's choice depends crucially on the size of the cost of monitoring. For a given value of the private benefit, there is a critical level of the cost of monitoring,

[16] These configurations of parameters satisfy Assumptions (1) and (2) and guarantee the existence of the equilibrium in both the single-bank and the two-bank game. The results do not change qualitatively with other parameter configurations. Note also that in Figure 6 the profit $\Pi^L_{F_1}$ depends on B only if $m > \overline{m}_1$, as the intensity M_1^* becomes an interior solution.

$\widetilde{m} \in (0, \overline{m}_1)$, such that $\Pi^L_{F_1} > \Pi^L_{F_2}$ if $m < \widetilde{m}$ and $\Pi^L_{F_1} > \Pi^L_{F_2}$ otherwise. If the cost of monitoring is very low, single-bank lending is optimal because it implies desirable high monitoring and low total monitoring costs. As m increases in the range $[0, \overline{m}_1]$, monitoring remains unaffected if the firm borrows from a single bank, whereas it falls if the firm borrows from two banks. This leads to lower expected financial return of the project, higher expected private return and eventually lower monitoring costs in (16) relative to (17). As m reaches the level $\widetilde{m}$, the last two effects dominate, and borrowing from two banks becomes optimal. To summarize:

Proposition 4 *The firm borrows from a single bank if the cost of monitoring is low, and from two banks otherwise.*

The firm prefers two-bank lending as a way of reducing overall monitoring and thus increasing its expected private return. Consequently, as Figure 6 shows, the threshold $\widetilde{m}$ decreases in the private benefit B. One can also show that the threshold $\widetilde{m}$ decreases in the success probability p_L. As long as $m < \overline{m}_1$, an increase in p_L reduces monitoring and elevates the firm's expected private return only if the firm borrows from two banks. Corollary 2 summarizes how the firm's optimal choice changes with the severity of its private and financial moral hazard problems.

Corollary 2 *The stronger the firm's private moral hazard and the weaker its financial moral hazard, the greater the set of values for the cost of monitoring for which two-bank lending is optimal.*

4 Extensions and Discussion of the Basic Model

In this section I analyze various modifications of the basic model, and discuss whether and in which fashion the previous results are affected.[17] Specifically, I address the case of linear monitoring costs, the availability of other non-bank sources of financing, and monitoring as a threat of liquidation.

4.1 Linear Monitoring Costs

Consider now that monitoring can be expanded at a constant marginal cost so that the cost function is linear and equal to $C(M) = \frac{m}{2}M$. The equilibrium monitoring intensities in the

[17] For reasons of brevity, I present and prove only the main new results. Detailed proofs of all results are available from the author upon request. Moreover, note that in Subsections 4.1 and 4.2 I take as given that misbehaving is the firm's optimal behavioral choice.

single-bank and the two-bank game are equal to:

$$M_{1L}^{*}=\begin{cases}1 & \textit{if } m\leq\overline{m}_{1L}=\frac{2(p_H-p_L)}{p_L}\\ \\ 0 & \textit{if } m>\overline{m}_{1L},\end{cases}\tag{18}$$

$$M_{2L}^{*}=\begin{cases}\frac{p_H-p_L(1+m)}{(p_H-p_L)(1+m)} & \textit{if } m\leq\overline{m}_{2L}=\frac{p_H-p_L}{p_L}\\ \\ 0 & \textit{if } m>\overline{m}_{2L}.\end{cases}\tag{19}$$

Expressions (18) and (19) show that, in contrast to the basic model, banks may choose not to monitor when the cost function is linear. Given $C'(0)>0$, there is no monitoring if the cost of the first marginal unit dominates its benefit. In this case the project is not financed, since, under Assumption (1), banks would not break even. If monitoring is optimal, the single bank always monitors more than the two banks in the aggregate, as in the basic model. However, single-bank lending now implies a lower cost of financing.

Proposition 5 *If the monitoring cost function is linear, the two banks always require a higher loan rate than the single bank.*

In the absence of diseconomies of scale in monitoring, the single bank is always more cost-effective than the two banks, since it has better monitoring incentives and avoids the duplication of effort.[18] Nevertheless, the firm may still want to borrow from two banks so as to reduce overall monitoring and enjoy a higher expected private return. The main intuition underlying Proposition 4 still holds. The firm prefers single-bank lending if the cost of monitoring is low, and it eventually prefers two-bank lending as m increases. The reason is that, even if the two banks become relatively more costly as m grows, their overall monitoring decreases and the firm's expected private return rises. As m and B increase sufficiently, these latter effects dominate, and two-bank lending becomes optimal.

4.2 Other Outside Investors

So far I have assumed that monitoring is not contractable and that only bank lending is available. In this context, the firm cannot influence a bank's monitoring incentive, and it prefers to borrow from two banks as a way to reduce overall monitoring if the cost of monitoring and the private benefit are sufficiently high. As an alternative way to reduce monitoring, suppose now that the firm can borrow from *both* a single bank and other dispersed outside investors, such as trade creditors (henceforth, "bank-investor game").[19] Let M_{1T} be the monitoring

[18] A similar idea is described by Winton (1993) in the context of shareholders' monitoring.

[19] Although the dispersed outside investors do not exert any monitoring, they correctly anticipate the monitoring intensity exerted by the bank in equilibrium.

intensity of the bank, and $r_{1T}+T$ be the total repayment of the bank and the outside investors. Proposition 6 characterizes the competitive equilibrium of the bank-investor game. Definition 3 introduces the function that will describe the equilibrium total repayment.

Definition 3 Let $g_{1T}(m)$ be the following function:

$$g_{1T}(m)=\begin{cases}\frac{2+m}{2p_H} & \textit{if } m\leq\overline{m}_{FB}=(p_H-p_L)R-B\\ \frac{((p_H-p_L)R-B)^2+2m}{2((p_H-p_L)^2R-(p_H-p_L)B+p_Lm)} & \textit{if } \overline{m}_{FB}<m<\overline{m}_{1T}=\frac{(p_H-p_L)^2R^2-B^2}{2(1-p_LR)}\\ R & \textit{if } m\geq\overline{m}_{1T}.\end{cases}$$

Proposition 6 *The unique competitive equilibrium of the bank-investor game in which the project is financed has the following features:*
1) there is a critical level of the cost of monitoring, $\overline{m}_{1T}$, such that the return of the project is $R=g_{1T}(m)$ if $m\geq\overline{m}_{1T}$, and $R<g_{1T}(m)$ otherwise;
2) the cost of monitoring satisfies the condition $m\leq\frac{(p_H-p_L)^2R^2}{2(1-p_LR)}$;
3) the total repayment of the bank and the outside investors is $r^*_{1T}+T^*=g_{1T}(m)$;
4) there is a critical level of the cost of monitoring, $\overline{m}_{FB}$, such that the bank monitors with intensity $M^*_{1T}=1$ if $m\leq\overline{m}_{FB}$, and with $M^*_{1T}<1$ otherwise. More precisely,

$$M^*_{1T}=\begin{cases}1 & \textit{if } m\leq\overline{m}_{FB}\\ \frac{(p_H-p_L)R-B}{m} & \textit{if } \overline{m}_{FB}<m<\overline{m}_{1T}\\ \frac{(p_H-p_L)R-\sqrt{(p_H-p_L)^2R^2-2m(1-p_LR)}}{m} & \textit{if } m\geq\overline{m}_{1T}.\end{cases}\tag{20}$$

The main result of Proposition 6 is that the single bank exerts a (weakly) lower monitoring intensity when it finances the firm with other investors than when it is the sole lender. In contrast to Proposition 1, the monitoring intensity M^*_{1T} is an interior solution over a greater set of values for the cost of monitoring, and is decreasing in the private benefit B. Moreover, M^*_{1T} equals the first-best level as long as $m<\overline{m}_{1T}$, where $\overline{m}_{1T}>\overline{m}_{FB}$. Thus, borrowing from a single bank and outside investors mitigates the overmonitoring problem of Proposition 1. The intuition is that, by choosing the amount of funds that it wants to borrow from the bank and by setting the repayment r^*_1 appropriately, the firm can affect the bank's monitoring incentive and induce it to exert the first-best monitoring level. This works, however, only as long as the feasibility constraint, $r^*_{1T}+T^*\leq R$, is not binding, i.e., as long as $m<\overline{m}_{1T}$. As the cost of monitoring increases beyond $\overline{m}_{1T}$, the firm must increase the amount it borrows from the bank and its repayment so as to induce the bank to exert a higher monitoring intensity. This increases the total success probability of the project, and makes external financing still feasible. Thus, as m increases beyond $\overline{m}_{1T}$, M^*_{1T} starts to increase with the cost of monitoring, and the total repayment of the bank and the outside investors equals the entire return of the project.[20]

Using numerical analysis once again, I compare the equilibrium of the bank-investor game with the one of the two-bank game, and I discuss whether and under which conditions two-bank lending is still optimal. Figure 7 illustrates the monitoring intensities in the bank-investor and the two-bank game as m increases in the interval $[0, 1]$. Figure 8 illustrates the total repayment in the bank-investor game and the loan rate in the two-bank game as m increases in the interval $[0, 1]$. The curves M_{1T}^* and $r_{1T}^*+T^*$ are drawn for $B = \{0.4, 0.5, 0.6\}$, and $p = 0.4$ in both figures.[21] Corollaries 3 and 4 summarize the main results of the simulations.

Figures 7 and 8 about here

Corollary 3 *The monitoring intensity of the single bank co-financing the firm with outside investors is lower than the overall monitoring intensity of the two banks if the cost of monitoring is sufficiently high. The stronger the firm's private moral hazard, the greater the set of values for the cost of monitoring for which this result obtains.*

Corollary 4 *The total repayment of the single bank and the outside investors is higher than the loan rate of the two banks if the cost of monitoring is sufficiently high. The stronger the firm's private moral hazard, the greater the set of values for the cost of monitoring for which this result obtains.*

Corollary 3 suggests that, if m is sufficiently high, borrowing from a single bank and outside investors is a better way to reduce monitoring than borrowing from two banks. The main reason is that the firm can affect the single bank's monitoring incentive, but it cannot influence the two banks' incentives. Thus, Corollary 3 also suggests that two-bank lending may imply an overmonitoring problem with respect to the first-best level. This problem emerges if the cost of monitoring and the private benefit are sufficiently high, since the duplication of effort limits the reduction in the two banks' overall monitoring as m increases, and the private benefit does not affect their monitoring decisions. Corollary 4 strengthens the result of Proposition 3 in that borrowing from two banks is cheaper than borrowing from a single bank and outside investors if the cost of monitoring and the private benefit are sufficiently high.

Consider now the firm's choice between borrowing from a single bank and outside investors (henceforth, "bank-investor" lending) and borrowing from two banks. Figures 9 and

[20] The result that the need for outside finance can lead to higher non-contractable effort is also in Aghion *et al.* (1999), where a firm can improve its incentive to exert effort by contracting on verifiable investments.

[21] As before, the other exogenous parameters in the simulations are $p_H = 1$ and $R = 1.6$.

10 illustrate the firm's expected profits in the two cases as m increases in the interval $[0, 1]$. Figure 9 is drawn for $p_L = 0.4$ and Figure 10 for $p_L = 0.6$.

Figures 9 and 10 about here

One can see that the firm's optimal choice may change sharply from the basic model as p_L and B increase. As Figure 9 illustrates, the results of the basic model still hold when $p_L = 0.4$. The firm borrows from a single bank and outside investors if the cost of monitoring is low, and it borrows from two banks as m increases sufficiently. As Figure 10 shows, however, when $p_L = 0.6$ and $B = 0.6$, the firm always borrows from a single bank and outside investors.

To understand the intuition behind this result, one has to again consider the trade-off in the firm's decision between monitoring and cost of financing. When misbehavior is accompanied by a high private benefit, the firm does not want the banks to exert much monitoring, and it would prefer the borrowing structure that minimizes it. Low monitoring, however, reduces the expected financial return of the project, thus increasing the cost of financing as a compensation for higher risk. The firm must balance the benefit of low monitoring against the drawback of high cost of financing, and it may not choose the monitoring-minimizing borrowing structure if it is too expensive. But an increase in p_L allows low monitoring without too high a compensation adjustment in the cost of financing, since it elevates the expected financial return of the project and reduces the value of monitoring. Thus, when both p_L and B are sufficiently high, the firm prefers to borrow from a single bank and outside investors. In practise, this suggests that when things are going well, it becomes easier for the firm to reduce monitoring and enjoy high private benefits.[22] Proposition 7 summarizes the firm's optimal choice.[23]

Proposition 7 *When the firm's private moral hazard is very strong and its financial moral hazard is weak, the firm always borrows from a single bank and outside investors. In all other circumstances, the results of Proposition 4 hold.*

[22] This seems to suggest a positive relationship between p_L and B. For example, when things go well, it may be relatively easier not to work and enjoy a quiet life. When firms are more profitable, accounting standards and transparency rules may be softer and leave room for larger private benefits. Nevertheless, the converse may also be true. For instance, riskier projects may entail higher private benefits. In this case, monitoring cannot be reduced much, and, as in the basic model, two-bank lending would dominate for a sufficiently high cost of monitoring and a sufficiently high private benefit.

[23] Given this result, one could argue that borrowing from two banks and outside investors would always dominate, as it reduces monitoring and provides the benefits of diseconomies of scale. This is not completely obvious. While convex monitoring costs would lead, all else equal, to a lower cost of funding, the duplication of effort would limit the reduction in monitoring. The net effect would be somewhat ambiguous.

4.3 Monitoring as a Threat of Liquidation

So far I have assumed that banks have a direct control on the firm, since monitoring allows them to observe the firm's behavior and to intervene in case of misbehavior. As an alternative form of control, consider now that monitoring enables banks to observe the firm's behavior before the project matures and to liquidate the firm for a total value of K.[24] Liquidation is preferable only if the bank observes misbehavior, i.e., $p_L R < K < p_H R$, and the firm always continues the project if there is no monitoring.[25] This alternative monitoring technology modifies the basic model as follows. First, it implies a different equilibrium of the model. The behavior/monitoring subgame of both the single-bank and the two-bank game now has only one mixed strategy equilibrium, in which the firm is indifferent between behaving and misbehaving and a bank chooses its monitoring intensity taking into account the firm's incentive to behave well. Second, monitoring improves the firm's incentive to behave well. Thus, the firm behaves better with single-bank lending than with two-bank lending, since the duplication of effort and the sharing of monitoring benefits reduce the two banks' monitoring incentives. Third, in contrast to Lemma 1, the single bank monitors as much as the two banks in the aggregate, when loan rates are assumed to be the same in the two games.

Despite these differences, the firm's optimal choice between borrowing from a single bank and two banks does not change substantially from that in the basic model. To see this, consider the firm's expected profits in equilibrium with single-bank and two-bank lending:

$$\Pi^L_{F_{1K}} = [e_1^* p_H + (1-e_1^*)(1-M_{1K}^*)p_L]\,R + (1-e_1^*)(1-M_{1K}^*)B - [1 + \frac{m}{2}(M_{1K}^*)^2], \quad (21)$$

$$\Pi^L_{F_{2K}} = [e_2^* p_H + (1-e_2^*)(1-\overline{M}_{2K}^*)p_L]R + (1-e_2^*)(1-\overline{M}_{2K}^*)B - [1 + m(M_{2K}^*)^2], \quad (22)$$

where e_1^* and e_2^* are the equilibrium probabilities with which the firm behaves well in the two scenarios, and the other variables have the usual meanings.

As in the basic model, the firm's choice depends on the relative differences between expected financial return of the project, expected private return and total monitoring costs in expressions (21) and (22). In contrast to the basic model, however, one must also take account now of how the firm's behavior changes with the parameters of the model.

Consider the cost of monitoring first. Single-bank lending is optimal if m is low, since it implies desirable high monitoring, improves the firm's behavior, and lowers total monitoring costs relative to two-bank lending. As m increases, the two banks reduce monitoring less than the single bank because the duplication of effort becomes smaller. As a consequence,

[24] The value K can be interpreted as the liquidation value of the firm or as collateral.

[25] The idea is similar to the one in Rajan and Winton (1995) and Park (2000). If a bank observes misbehavior through monitoring, it can force the firm into default by invoking a covenant, or by refusing to refinance a short-term loan.

the decline in e_2^* is less than the decline in e_1^*. Along with larger diseconomies of scale and lower total monitoring costs, this effect eventually makes two-bank lending more attractive than single-bank lending.

Consider now the private benefit. An increase in B elevates monitoring and worsens the firm's incentive to behave well. The two banks increase monitoring less than a single bank; consequently, e_2^* reduces now more than e_1^*. This makes single-bank lending more attractive relative to two-bank lending, since it increases the expected financial return of the project in (21) by more than in (22). On the other hand, however, increasing B elevates the firm's expected private return in (22) by more than in (21) because $e_2^* < e_1^*$. This effect makes two-bank lending relatively more attractive, and it dominates as the cost of monitoring and the private benefit increase sufficiently. Similar arguments hold for a change in the success probability of the project p_L.

To sum up, the predictions of the basic model concerning the optimal borrowing structure do not change if monitoring leads to liquidation of the project rather than to a direct control over the firm's behavior. The main idea is that two-bank lending may still be optimal as a way to reduce overall monitoring and the firm's incentive to behave well. In both approaches, two-bank lending involves an externality in that a bank's monitoring induces good behavior on the whole project in the basic model and it leads to sharing of the value K in case of monitoring as threat of liquidation. Removing this externality improves the incentives of both the two banks and the firm, but it may not be optimal. To see this, suppose that a monitoring bank is the first to seize the value K in case of liquidation. This reduces the free-riding problem of the two banks, and increases both the monitoring intensity M_{2K}^* and the firm's effort e_2^*. In terms of expression (22), these effects increase the expected financial return of the project, but they also lower the expected private return and increase total monitoring costs. If the latter two effects dominate, reducing the free-rider problem of the two banks lowers the firm's expected profit, and two-bank lending may no longer be optimal.

5 Concluding Remarks

Monitoring of borrowers is undoubtedly one of the key functions that banks perform in the economy. Yet the literature has not paid much attention to the question of how much monitoring banks exert and to the effects of this on the cost of financing, the quality of borrowers and the number of banks a firm should borrow from. This paper analyzes a firm's choice between borrowing from a single bank and two banks in a context where the number of banks influences both the level of monitoring and loan rates. The results of the model–which are robust to different specification of the monitoring technology–are particularly applicable to the financing of small businesses, given the need for monitoring in granting loans and the lack of public information about such firms.

The number of banks affects monitoring and loan rates in various ways. Borrowing from two banks involves duplication of effort and sharing of monitoring benefits. This implies that a bank monitors more when it is the sole lender than when it finances the firm with another bank. Yet borrowing from two banks does not necessarily imply a higher cost of financing. If there are diseconomies of scale in monitoring (reflecting the increasing difficulty of discovering more about a firm one already knows a lot about, or scarce resources for adequate monitoring), two-bank lending is cheaper than single-bank lending whenever the technological effect of the convex monitoring cost function dominates the duplication of effort and the sharing of monitoring benefits. In contrast to the literature based on the hold-up and the soft-budget-constraint problems, these results provide a theoretical rationale for the empirical observation that increasing the number of banks tends to lower firms' quality, while either increasing or decreasing the cost of financing (e.g., Petersen and Rajan, 1994; Foglia *et al.*, 1998; D'Auria *et al.*, 1999).

The firm's choice between single-bank and two-bank lending depends on how both monitoring and loan rates differ in the two scenarios. The optimal choice balances the benefit of monitoring in terms of higher expected financial return of the project against its drawbacks in terms of lower expected private return and higher total monitoring costs. The attractiveness of two-bank lending is increasing in the cost of monitoring, the private benefit, and the firm's expected profitability.

The model has a number of empirical implications. First, the cost of monitoring relates to the ease with which banks are able to obtain and process information about firms. Factors affecting such a cost include disclosure rules, accounting standards, and, more generally, legal and regulatory aspects of the financial system, like the degree of investor protection, the efficiency of the judicial system, and legal enforcement. Similar elements contribute to the ability with which entrepreneurs are able to obtain and enjoy private benefits, since they limit the discretionary power of management (La Porta *et al.*, 1997, 1998; Dyck and Zingales, 2001).[26] Thus, the model predicts that firms should prefer multiple-bank lending in countries with laxer accounting and disclosure standards, lower investor protection, and more inefficient judicial systems. Second, the model predicts that firms with higher expected profitability should prefer multiple-bank lending. These predictions are supported by Ongena and Smith (2000b) and Detragiache et al. (2000), who find greater use of multiple-bank lending in countries with weaker judicial systems and weaker investor protection. Moreover,

[26] Dyck and Zingales (2001) find that other extra-legal variables, such as a high diffusion of the press and a high rate of tax compliance, also reduce private benefits.

the predictions of the model are consistent with the findings in Dyck and Zingales (2001) of high private benefits in countries like Italy and Portugal, where multiple-bank lending is used extensively.[27]

A last important insight of the model is the potential role of other sources of financing. The analysis suggests that firms with profitable projects and high private benefits prefer to borrow from a single bank and other dispersed investors as a way to reduce monitoring, without too high a compensating adjustment in the cost of financing.

The paper proposes an alternative to the hold-up and soft-budget-constraint problems in addressing the question of why multiple-bank lending can be optimal despite involving duplication of effort and sharing of monitoring benefits. In this respect, focusing on a one-period model and assuming a competitive banking system are strengths of the model. Extending the time structure and allowing for an ex post imperfect banking sector would simply make a stronger case for multiple-bank lending. This analysis, however, is left for future research.

Appendix

Proof of Proposition 1

I solve the game backwards. First, I take the loan rate r_1 as given, and I find the Nash equilibrium of the behavior/monitoring subgame; second, I solve the loan rate setting in the first stage.

i) Behavior/monitoring subgame. I look for the pure strategy Nash equilibrium of the behavior/monitoring subgame; it can be shown that this is without loss of generality (see Carletti, 2001). If the firm chooses H, the bank maximizes (5), and chooses $M_1^*(r_1) = 0$. This is not an equilibrium. If $M_1^*(r_1) = 0$, the firm deviates and chooses L, since under Assumption (1) (4)>(3) for all $M_1(r_1) < 1$. If the firm chooses L, the bank chooses $M_1(r_1)$ so as to maximize (6), which is differentiable and strictly concave in $M_1(r_1)$. The first-order condition gives

$$\frac{\partial \Pi_{B_1}^L}{\partial M_1(r_1)} = (p_H - p_L)r_1 - mM_1(r_1) = 0.$$

Since $M_1(r_1)$ must be in $[0,1]$, $M_1^*(r_1) = \min\{\frac{(p_H-p_L)r_1}{m}, 1\}$. If $M_1^*(r_1) < 1$, the firm still prefers L; if $M_1^*(r_1) = 1$, it is indifferent between H and L. Therefore, the vector $(L, M_1^*(r_1))$ is the unique Nash equilibrium of the behavior/monitoring subgame.

[27] In a study based on duration analysis, Farinha and Santos (2002) find that firms with greater growth opportunities are more likely to substitute a single-bank relationship with multiple-bank relationships as the duration of that relationship increases. This finding seems to support my prediction of a positive relationship between firms' expected profitability and two-bank lending, even if the model focuses on the choice between single-bank and two-bank lending rather than on the switching between the two regimes.

ii) Loan rate setting. Setting equation (6) equal to zero after substituting $M_1^*(r_1)$ gives the equilibrium loan rate r_1^*. As $\Pi_{B_1}^L$ is continuous and increasing in r_1, there is only one value of r_1^* satisfying $\Pi_{B_1}^L = 0$, which coincides with $g_1(m)$. Substituting r_1^* in $M_1^*(r_1^*)$ gives (7). The equilibrium monitoring intensity is a corner solution if and only if $\frac{(p_H - p_L)r_1^*}{m} \geq 1$, which holds if and only if $m \leq \overline{m}_1$. The intensity M_1^* is an interior solution if and only if $\frac{(p_H - p_L)r_1^*}{m} < 1$, which holds if and only if $m > \overline{m}_1$. If $r_1^* \leq R$, the contract is feasible and the project is undertaken. Q.E.D.

Proof of Corollary 1

Substituting $r_1^* = g_1(m)$ in the feasibility condition $r_1^* \leq R$ implies (i) $m \leq 2(p_H R - 1)$ if $m \leq \overline{m}_1$, and (ii) $m \leq \frac{(p_H - p_L)^2 R^2}{2(1 - p_L R)}$ if $m > \overline{m}_1$. For (ii) to be possible, it must hold $\frac{(p_H - p_L)^2 R^2}{2(1 - p_L R)} \geq \overline{m}_1$. This requires $R > \frac{2}{p_H + p_L}$, which implies $2(p_H R - 1) > \overline{m}_1$. Thus, if $R > \frac{2}{p_H + p_L}$, condition (i) is always satisfied, and the statement follows. Q.E.D.

Proof of Proposition 2

As for Proposition 1, the game is solved backwards. I start with the Nash equilibrium of the behavior/monitoring subgame; then I solve the loan rate setting in the first stage.
i) Behavior/monitoring subgame. I look only for the pure strategy Nash equilibrium of the behavior/monitoring subgame; it can be shown that there does not exist any mixed strategy equilibrium (see Carletti, 2001). With a similar argument to that in the proof of proposition 1, it can be easily shown that the firm chooses L. Given this, if bank B chooses $M_B(r_2)$, bank A chooses $M_A(r_2)$ so as to maximize (12), which is continuous and strictly concave in $M_A(r_2)$. The first-order condition gives

$$\frac{\partial \Pi_{B_A}^L}{\partial M_A(r_2)} = (1 - M_B(r_2))(p_H - p_L)\frac{r_2}{2} - mM_A(r_2) = 0,$$

from which it is

$$M_A(r_2) = \frac{(p_H - p_L)r_2}{2m}(1 - M_B(r_2)).$$

Substituting $M_A(r_2) = M_B(r_2)$ in a symmetric equilibrium gives $M_2^*(r_2) = M_i^*(r_2)$ for $\forall i = A, B$.
ii) Loan rate setting. Setting equation (12) equal to zero after substituting $M_i^*(r_2)$ gives the equilibrium loan rate r_2^*. The expression $\Pi_{B_i}^L = 0$ coincides with $f(g_2(m)) = 0$. Since the function $f(g_2)$ is cubic in g_2, with a positive coefficient on g_2^3, $f(0) = -4m^2 < 0$ and $f'(0) < 0$, there is only a unique real positive solution $r_2^* = g_2(m)$. Substituting r_2^* in $M_i^*(r_2^*)$ gives (13). If $r_2^* \leq R$, the contract is feasible and the project is undertaken. Q.E.D.

Proof of Lemma 1

Define $diff_M = M_1^*(r) - \overline{M}_2^*(r)$. Substituting the expressions for $M_1^*(r)$ and $\overline{M}_2^*(r)$ when $r_1 = r_2 = r$ gives: (i) if $m \leq (p_H - p_L)r$, $diff_M > 0$ as $M_1^*(r) = 1$ and $\overline{M}_2^*(r_2) < 1$; (ii) if $m > (p_H - p_L)r$, $diff_M = \frac{((p_H - p_L)r)^2(3m + (p_H - p_L)r)}{m((p_H - p_L)r + 2m)^2} > 0$. Q.E.D.

Proof of Lemma 2

Define $diff_C = C(M_1^*(r)) - 2C(M_2^*(r))$. Substituting the expressions for $C(M_1^*(r_1))$ and $2C(M_2^*(r_2))$ when $r_1 = r_2 = r$ gives: (i) if $m \leq (p_H - p_L)r$, $diff_C = \frac{m}{2} - m(\frac{(p_H - p_L)r}{(p_H - p_L)r + 2m})^2$ which is positive only if $m > \frac{(p_H - p_L)r}{2(1+\sqrt{2})}$; (ii) if $m > (p_H - p_L)r$, $diff_C = \frac{(p_H - p_L)^2 r^2}{2m} - m(\frac{(p_H - p_L)r}{(p_H - p_L)r + 2m})^2 > 0$. Q.E.D.

Proof of Proposition 5

Solving the relevant expressions for banks' profits equal to zero after substituting (18) and (19) gives the equilibrium loan rates $r_{1L}^* = \frac{2+m}{2p_H}$ and $r_{2L}^* = \frac{1+m}{p_H}$. The result follows immediately. Q.E.D.

Proof of Proposition 6

Let I and E be the amounts that the firm borrows from the bank and other outside investors. The competitive equilibrium is the solution to the following problem:

$$\underset{I, r_{1T}, E, T}{Max} \ \Pi_{F_{1T}}^L = [M_{1T}^*(r_{1T})p_H + (1 - M_{1T}^*(r_{1T}))p_L](R - r_{1T} - T) + (1 - M_{1T}^*(r_{1T}))B \quad (23)$$

subject to

$$M_{1T}^*(r_{1T}) = \frac{(p_H - p_L)r_{1T}}{m} \quad (24)$$

$$[M_{1T}^*(r_{1T})p_H + (1 - M_{1T}^*(r_{1T}))p_L]r_{1T} = I + \frac{m}{2}(M_{1T}^*(r_{1T}))^2 \quad (25)$$

$$[M_{1T}^*(r_{1T})p_H + (1 - M_{1T}^*(r_{1T}))p_L]T = E \quad (26)$$

$$r_{1T} + T \leq R \quad (27)$$

where (23) is the firm's expected profit, (24) is the bank's incentive compatibility constraint, that is its profit-maximizing monitoring level, equations (25) and (26) are the zero-profit conditions for the bank and the outside investors, and (27) is the contract feasibility condition.

Substituting (24), (25) and (26) in (23)and (27), and using $E = 1 - I$ allows the problem to be restated only in terms of r_{1T}. The constraint (27) then reduces to

$$\frac{1 + \frac{m}{2}(\frac{(p_H - p_L)r_{1T}}{m})^2}{[\frac{(p_H - p_L)r_{1T}}{m}p_H + (1 - \frac{(p_H - p_L)r_{1T}}{m})p_L]} \leq R. \quad (28)$$

Denoting as λ the Kuhn-Tucker multiplier associated to the constraint (28), the necessary and sufficient condition for the optimal r_{1T} is given by

$$\frac{(p_H - p_L)^2 R}{m} - \frac{(p_H - p_L)^2 r_{1T}}{m} - \frac{(p_H - p_L)B}{m} + \lambda[p_L \frac{(p_H - p_L) r_{1T}}{m} - \frac{(p_H - p_L)^2}{m}] = 0. \quad (29)$$

Suppose now that (28) is not binding, so that $\lambda = 0$. Equation (29) reduces to

$$r_{1T}^* = \frac{(p_H - p_L)R - B}{(p_H - p_L)}. \quad (30)$$

Substituting (30) in (24) and then in (28) implies statement 1. When (28) is binding, r_{1T}^* is given by

$$r_{1T}^* = \frac{(p_H - p_L)R - \sqrt{(p_H - p_L)^2 R^2 - 2m(1 - p_L R)}}{(p_H - p_L)}. \quad (31)$$

The expression for r_{1T}^* in (31) is a positive real solution only if $m \leq \frac{(p_H - p_L)^2 R^2}{2m(1 - p_L R)}$. This implies statement 2. Solving (26) after substituting (30) and (31) gives statement 3. Substituting then (30) and (31) in (24) implies statement 4. Q.E.D.

References

Aghion, P., Dewatripont, M., and Rey, P. (1999). Agency costs, firm behavior and the nature of competition, CEPR discussion papers no. 2130.

Besanko, D., and Kanatas, G. (1993). Credit market equilibrium with bank monitoring and moral hazard, *Rev. Finan. Stud.* **6**, 213-232.

Bolton, P., and Scharfstein, D. (1996). Optimal debt structure and the number of creditors, *J. Polit. Econ.* **104**, 1-25.

Boot, A.W. (2000). Relationship banking: What do we know?, *J. Finan. Intermed.* **9**, 7-25.

Burkart, M., Gromb, D., and Panunzi, F. (1997). Large shareholders, monitoring, and the value of the firm, *Quart. J. Econ.* **112**, 693-728.

Carletti, E. (2001). The structure of bank relationships, endogenous monitoring and loan rates, FMG discussion papers no. 388.

Cerasi, V., and Daltung, S. (2000). The optimal size of a bank: Costs and benefits of diversification, *Eur. Econo. Rev.* **44**, 1701-1726.

Covitz, D., and Heitfield, E. (1999). Monitoring, moral hazard and market power: A model of bank lending, FEDS working papers no. 99-37.

Degryse, H., and Ongena, S. (2001). Bank relationships and firm profitability, *Finan. Management* **30**, 9-34.

Detragiache, E., Garella, P., and Guiso, L. (2000). Multiple vs. single banking relationships: Theory and evidence, *J. Finance* **55**, 1133-1161.

Dewatripont, M., and Maskin, E. (1995). Credit and efficiency in centralized and decentralized economies, *Rev. Econ. Stud.* **62**, 541-555.

D'Auria, C., Foglia, A., and Marullo Reedtz, P. (1999). Bank interest rates and credit relationships in Italy, *J. Banking Finance* **23**, 1067-1093.

Dyck, A., and Zingales, L. (2001). Private benefits of control: An international comparison, Harvard Business School and University of Chicago, unpublished manuscript.

Elsas, R., and Krahnen, J.P. (1998). Is relationship banking special? Evidence from credit-file data in Germany, *J. Banking Finance* **22**, 1283-1316.

Farinha, L.A., and Santos, J.A. (2002). Switching from single to multiple bank lending relationships: Determinants and Implications, *J. Finan. Intermed.* **11**, 124-151.

Foglia, A., Laviola, S., and Marullo Reedtz, P. (1998). Multiple banking relationships and the fragility of corporate borrowers, *J. Banking Finance* **22**, 1441-1456.

Gorton, G., and Winton, A. (2003). Financial intermediation, *in* "Handbook of the Economics and Finance" (G. Constantinides , M. Harris and R. Stulz, Eds.), North Holland, Amsterdam, forthcoming.

Harhoff, D., and Körting, T. (1998). Lending relationships in Germany: Empirical evidence from survey data, *J. Banking Finance* **22**, 1317-1353.

Holmström, B., and Tirole, J. (1997). Financial intermediation, loanable funds, and the real sector, *Quart. J. of Econ.* **112**, 663-691.

La Porta, R., Lopez-de-Salines, F., and Shleifer, A. (1997). Legal determinants of external finance, *J. Finance* **53**, 1131-1150.

La Porta, R., Lopez-de-Salines, F., and Shleifer, A. (1998). Law and finance, *J. Polit. Econ.* **106**, 1113-1155.

Ongena, S., and Smith, D.C. (2000a). Bank relationships: A review, *in* “The Performance of Financial Institutions” (P. Harker and S.A. Zenios, Eds.), pp. 221-258, Cambridge University Press, Cambridge.

Ongena, S., and Smith, D.C. (2000b). What determines the number of bank relationships? Cross-Country evidence, *J. Finan. Intermed.* **9**, 26-56.

Padilla, A.J., and Pagano, M. (1997). Endogenous communication among lenders and entrepreneurial incentives, *Rev. Finan. Stud.* **10**, 205-236.

Pagano, M., and Röell, A. (1998). The choice of stock ownership structure: Agency costs, monitoring and the decision to go public, *Quart. J. Econ.* **113**, 187-225.

Park, C. (2000). Monitoring and the structure of debt contracts, *J. Finance* **55**, 2157-2195.

Petersen, M., and Rajan, R.G., (1994). The benefit of lending relationships: Evidence from small business data, *J. Finance* **49**, 1367-1400.

Rajan, R. (1992). Insiders and outsiders: The choice between informed and arm’s-length debt, *J. Finance* **47**, 1367-1400.

Rajan, R., and Winton, A. (1995). Covenants and collateral as incentives to monitor”, *J. Finance* **50**, 1113-1146.

Sharpe, S. (1990). Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships, *J. Finance* **45**, 1069-1087.

Thakor A. (1996). Capital requirements, monetary policy, and aggregate bank lending: Theory and empirical evidence, *J. Finance* **51**, 279-324.

Von Thadden, E.L. (1992). The commitment of finance, duplicated monitoring and the investment horizon, CEPR discussion papers no. 27.

Winton, A. (1993). Limitation of liability and the ownership structure of the firm, *J. Finance* **48**, 487-512.

Yosha, O. (1995). Information disclosure costs and the choice of financing source, *J. Finan. Intermed.* **4**, 3-20.

Figure 2: Monitoring intensities in the single-bank and the two-bank game. The figure shows how the monitoring intensity $M_1^*(r)$ of the single bank, the overall monitoring intensity $\overline{M}_2^*(r)$ and the individual monitoring intensity $M_2^*(r)$ of the two banks change as the cost of monitoring m increases. The loan rate is assumed to be the same in the two games, i.e., $r_1 = r_2 = r$.

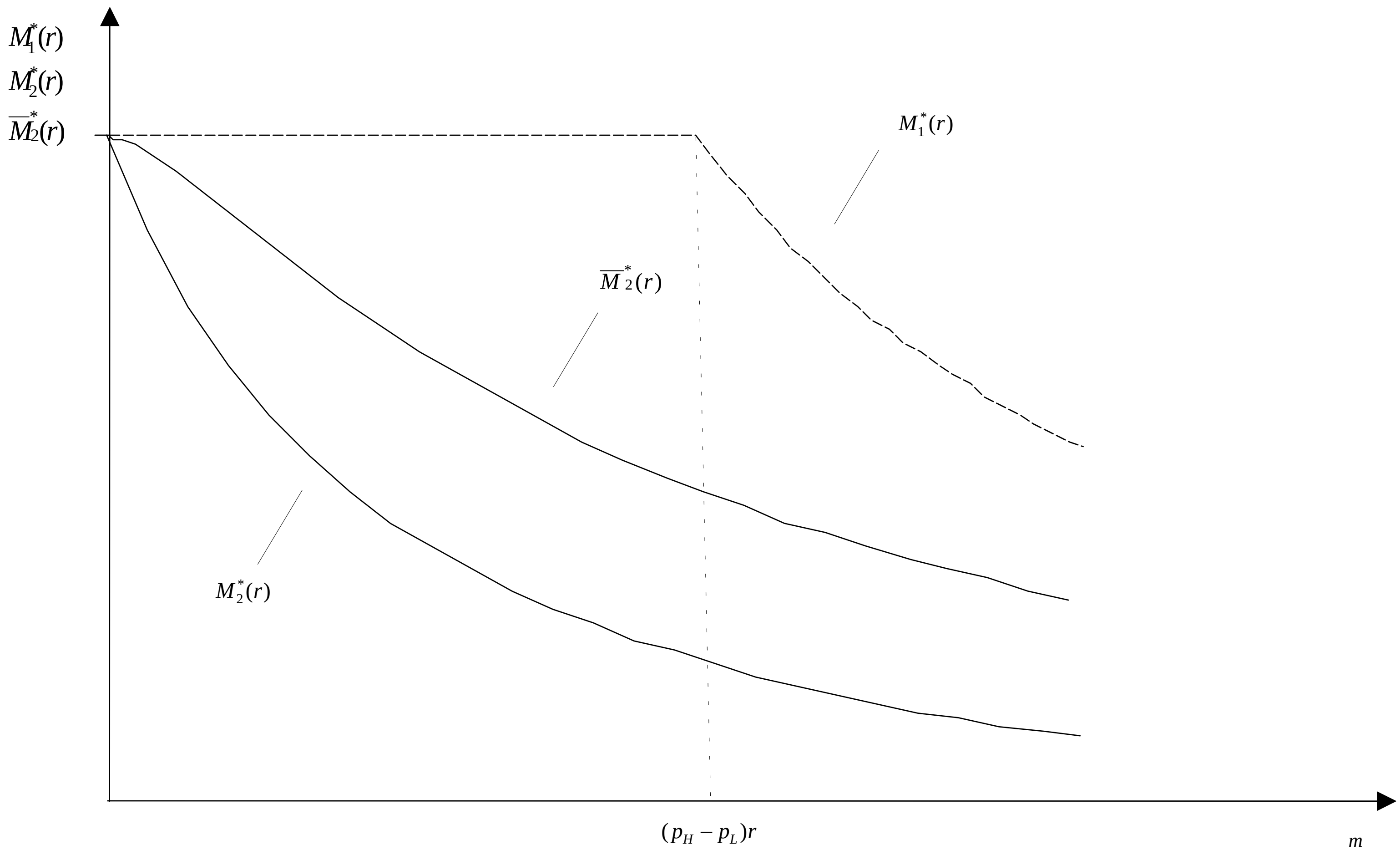

Figure 3: Monitoring costs in the single-bank and the two-bank game. The figure shows how the monitoring costs $C(M_1^*(r))$ of the single bank and the total monitoring costs $2C(M_2^*(r))$ of the two banks change as the cost of monitoring m increases. The loan rate is assumed to be the same in the two games, i.e., $r_1 = r_2 = r$.

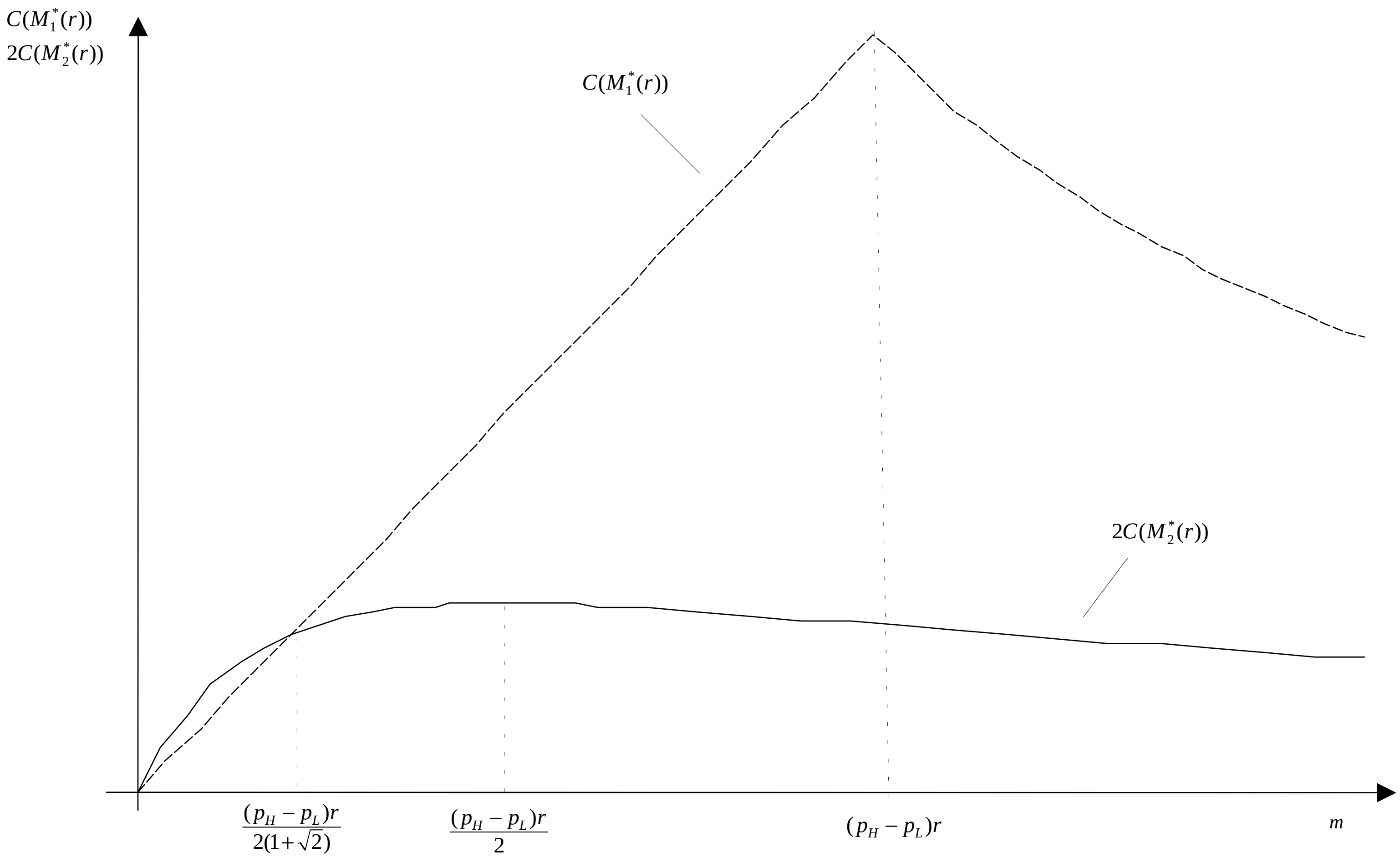

Figure 4: Loan rates in the single-bank and the two-bank game for a low success probability of the project. The figure shows how the loan rates r_1^* of the single bank and r_2^* of the two banks change as the cost of monitoring m increases. The figure is drawn for a success probability p_L = 0.4.

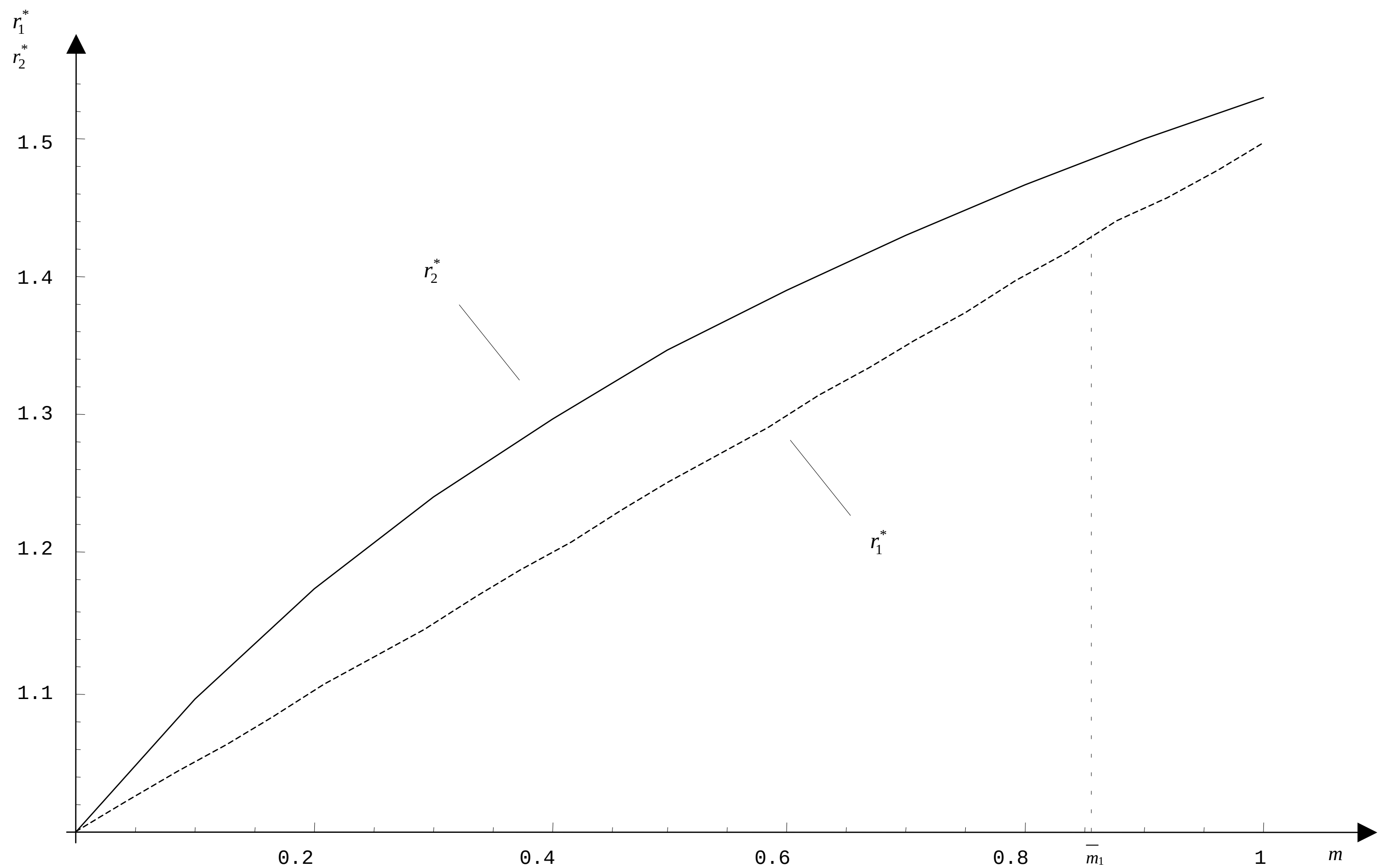

Figure 5: Loan rates in the single-bank and the two-bank game for a high success probability of the project. The figure shows how the loan rates r_1^* of the single bank and r_2^* of the two banks change as the cost of monitoring m increases. The figure is drawn for a success probability $p_L = 0.6$

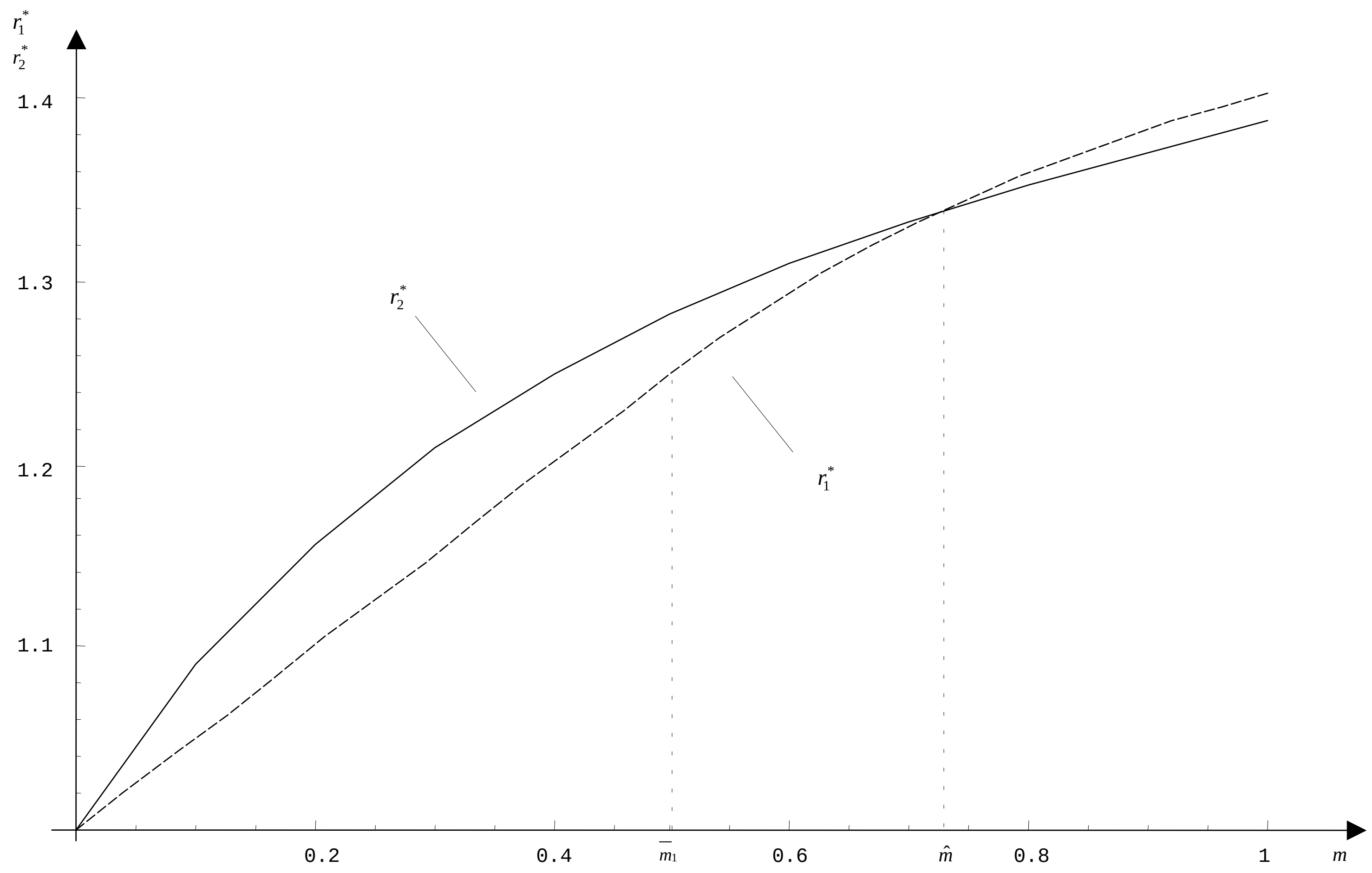

Figure 6: Firm's expected profits with single-bank and two-bank lending. The figure shows how the firm's expected profits $\Pi^L_{F_1}$ when it borrows from a single bank and $\Pi^L_{F_2}$ when it borrows from two banks change as the cost of monitoring m and the private benefit B increase. The figure is drawn for a success probability of the project $p_L = 0.4$.

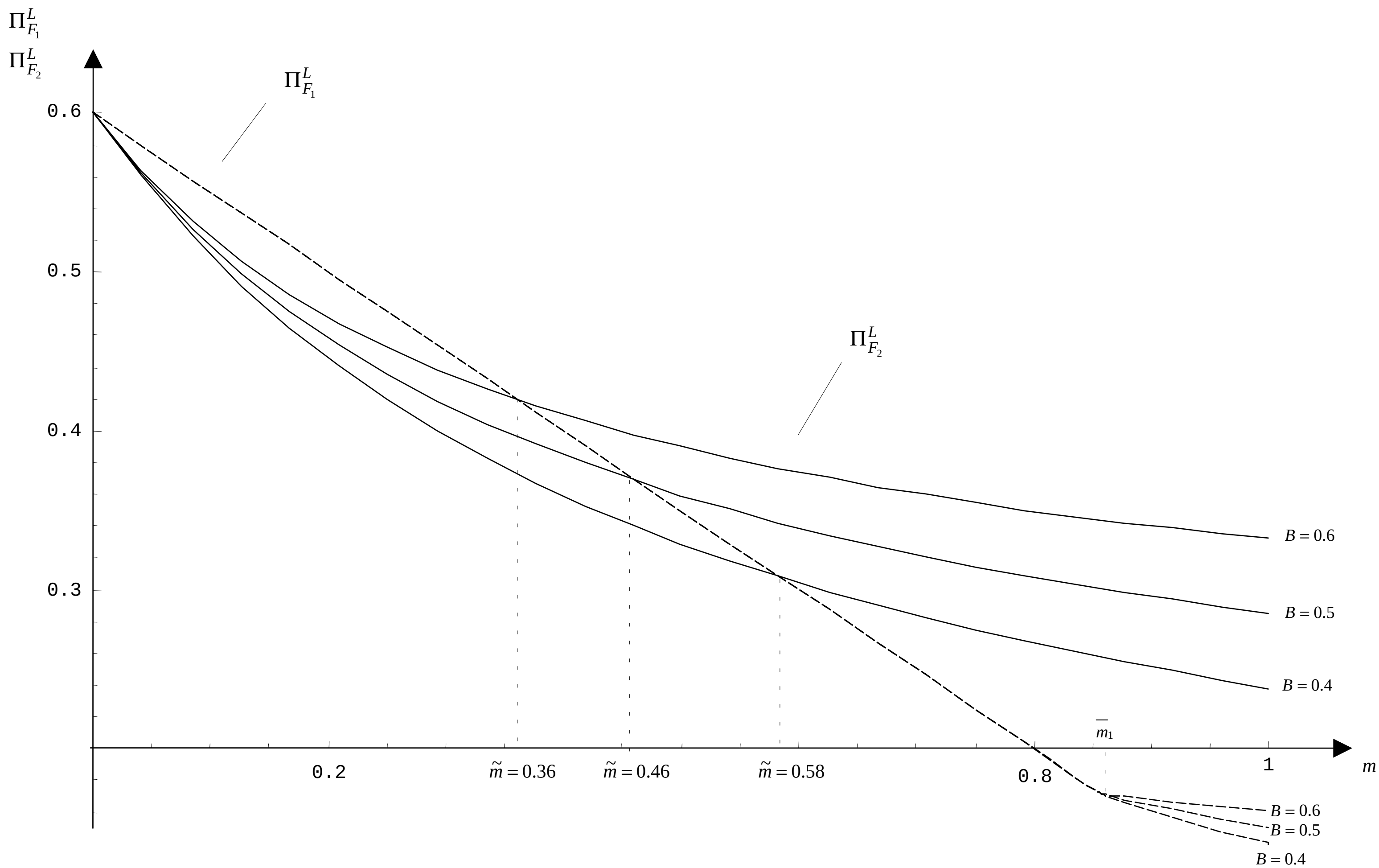

Figure 7: Monitoring intensities in the bank-investor and the two-bank game. The figure shows how the monitoring intensity M_{1T}^* of the single bank co-financing the firm with outside investors, the overall monitoring intensity $\overline{M}_2^*$ and the individual intensity M_2^* of the two banks change as the cost of monitoring m and the private benefit B increase. The figure is drawn for a success probability of the project $p_L = 0.4$.

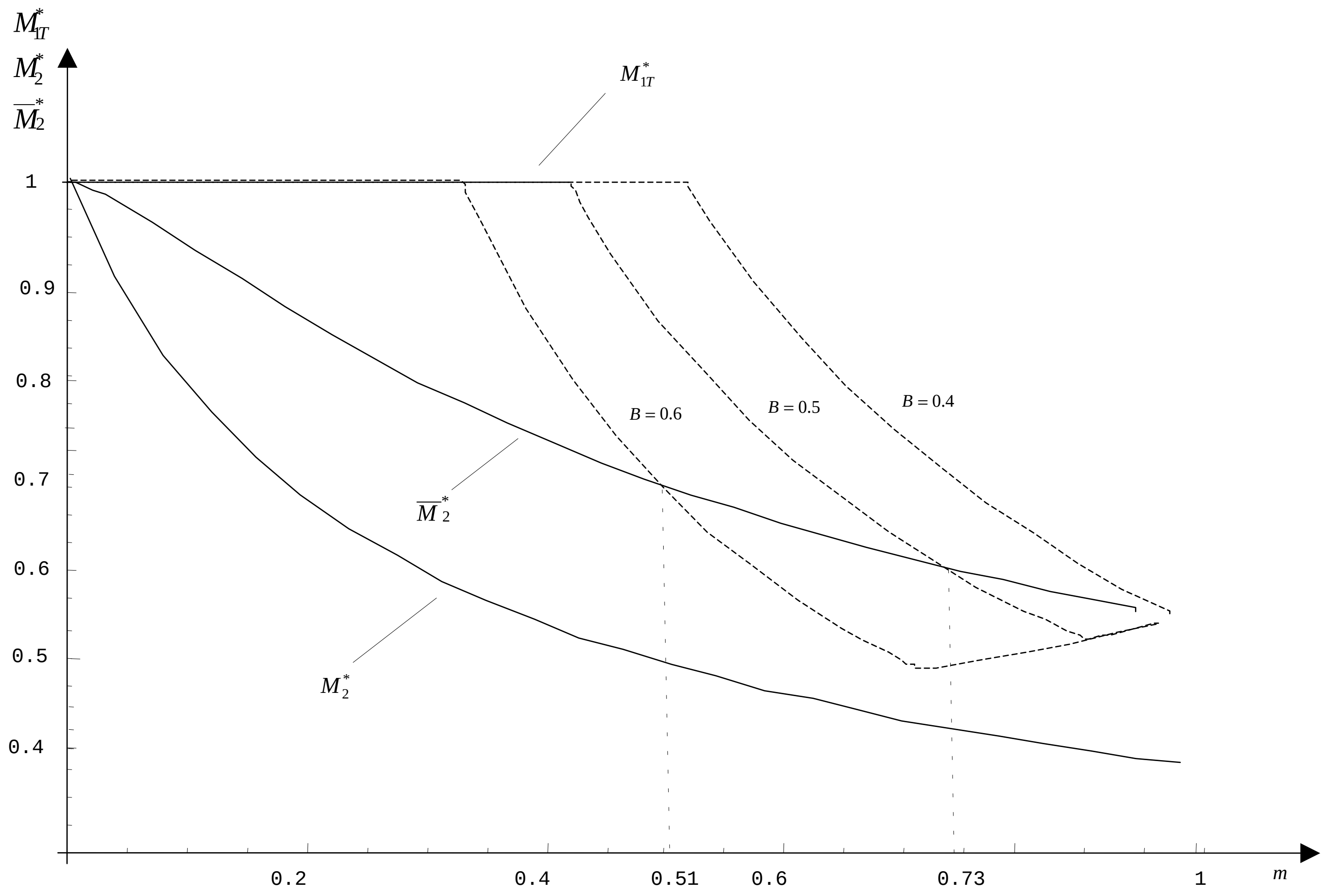

Figure 8: Total repayment in the bank-investor game and loan rate in the two-bank game. The figure shows how the total repayment $r_{1T}^{*}+T^{*}$ of the single bank and the outside investors and the loan rate r_2^{*} of the two banks change as the cost of monitoring m and the private benefit B increase. The figure is drawn for a success probability of the project $p_L = 0.4$.

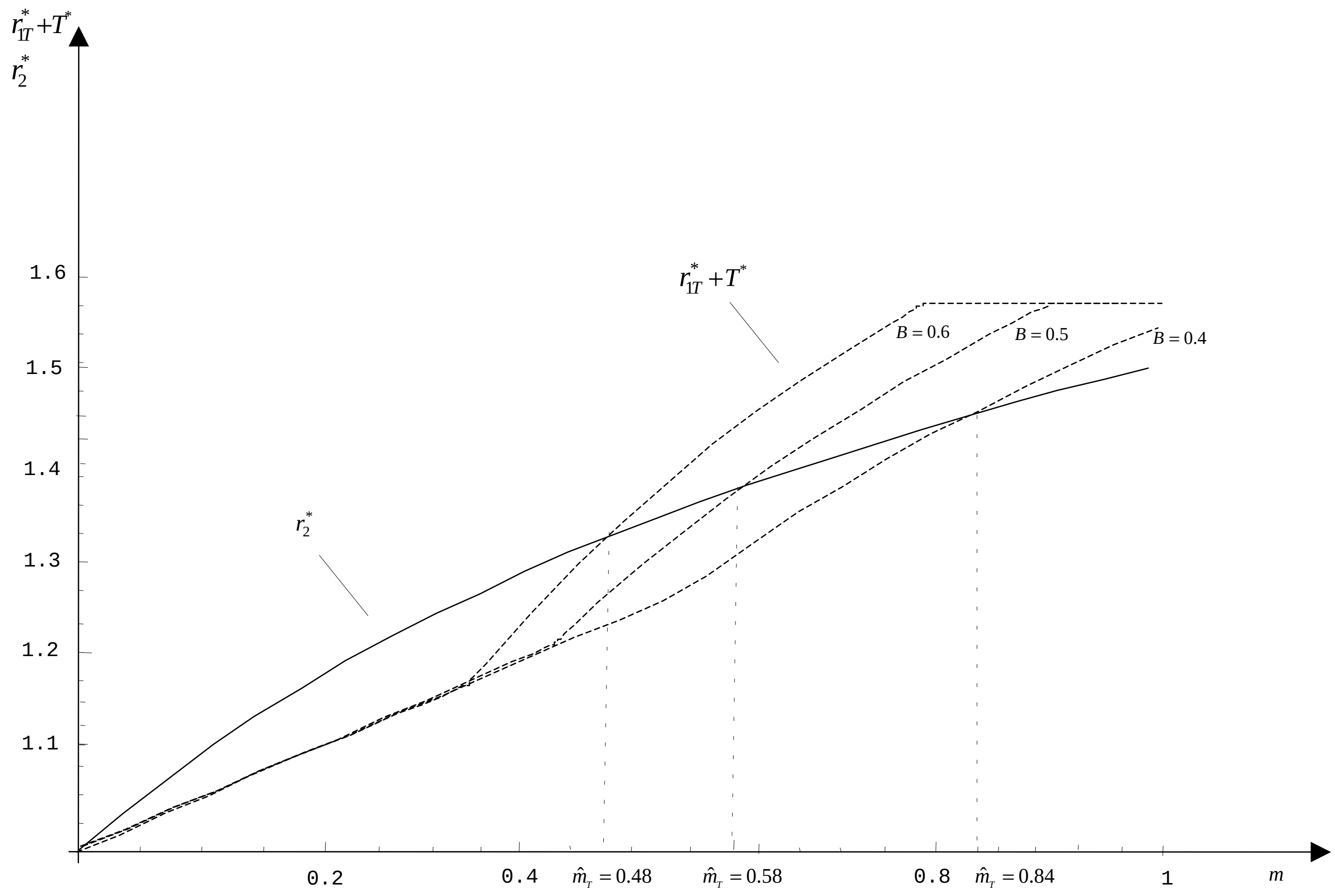

Figure 9: Firm's expected profits with bank-investor and two-bank lending for a low success probability of the project. The figure shows how the firm's expected profits $\Pi^L_{F_{1T}}$ when it borrows from a single bank and outside investors and $\Pi^L_{F_2}$ when it borrows from two banks change as the cost of monitoring m and the private benefit B increase. The figure is drawn for a success probability $p_L = 0.4$.

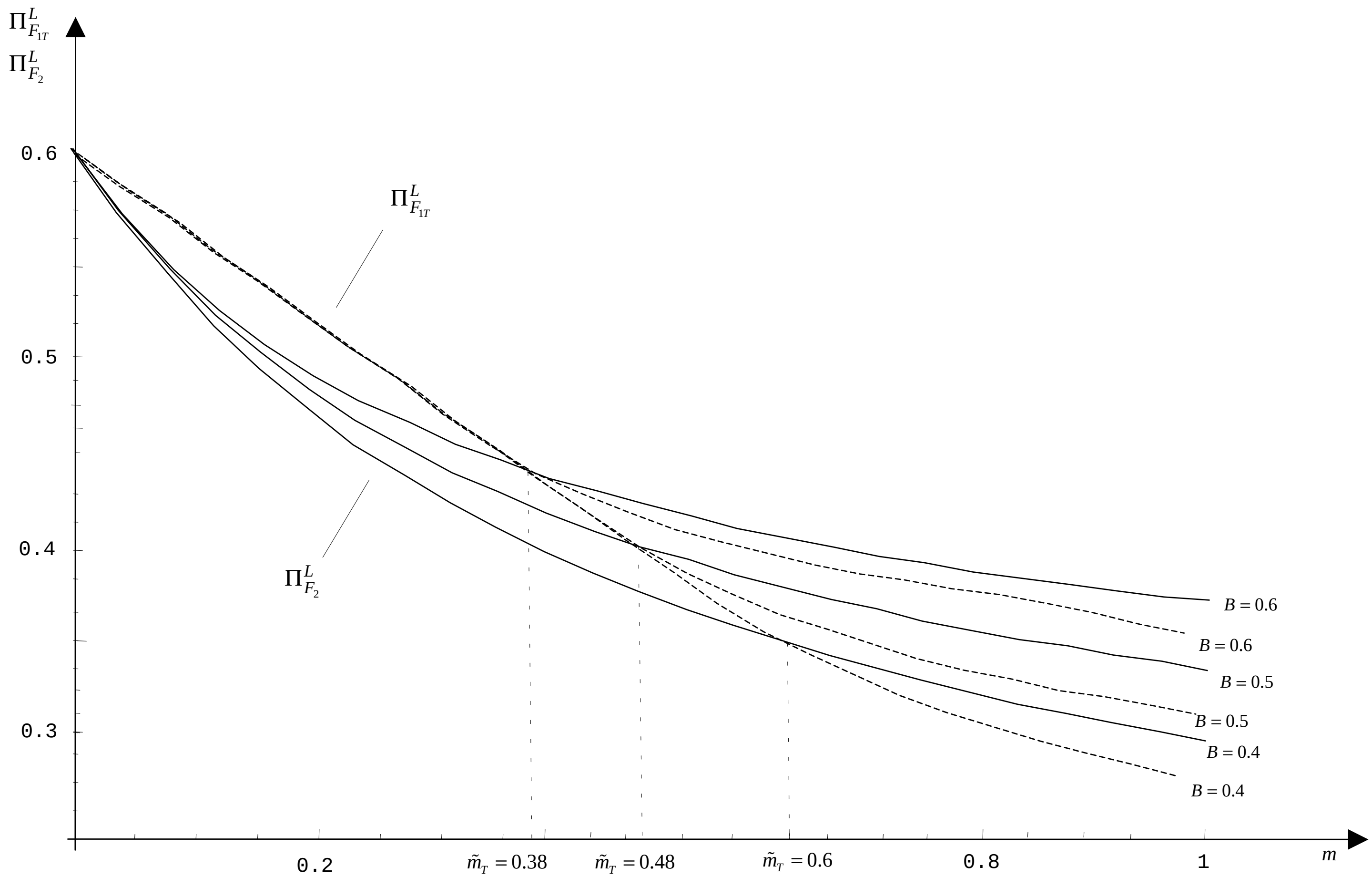

Figure 10: Firm's expected profits with bank-investor and two-bank lending for a high success probability of the project. The figure shows how the firm's expected profits $\Pi^L_{F_{1T}}$ when it borrows from a single bank and outside investors and $\Pi^L_{F_2}$ when it borrows from two banks change as the cost of monitoring m and the private benefit B increase. The figure is drawn for a success probability p_L = 0.6.