Investment, Perception of Risk and Financial Constraints

Alfonso Ugarte Ruiz

TESI DOCTORAL UPF / 2010

DIRECTOR DE LA TESI

Dr. Andrea Caggese (Departament d'Economia i Empresa)

UNIVERSITAT POMPEU FABRA

To Sofia and Sofia and the rest of my Family

Acknowledgements

I would like to specially thank my advisor Andrea Caggese for his guidance, support and suggestions during this long process of writing this thesis. He was always committed and ready to help and every advice I received was very useful to improve the thesis step by step. I would also like to thank all the people in UPF who have helped at some point with helpful comments, suggestions or encouragement, like Michael Greenacre, Vicente Cuñat, Xavier Freixas, Francisco Peñaranda, Joel Shapiro, Julian Rode, Oskar Nupia, Francesco Turino, Gueorgui Kolev, Stephan Litschig, Judit Montoriol, Ander Perez Orive, Francesco Caprioli and Alberto Martin among others. I would also like to thank Marta Araque and the rest of the staff at GPEFM for their invaluable support and constant help.

I would also like to specially thank those who helped me when I most needed: Elsa Ugarte, Amaya Ugarte, Luis Carreras, Sofía Ruiz and Julian Friend.

I also want to thank those of my friends who have been there encouraging me and helping in different ways in different times during these years: Leonardo Tariffi, Julian, Joro, Liliana Mazilu, Javier Alonso, David Freeman, Diego Lineros, Karina López and Ivan Piñerez.

Last but not least, my family again, my mother, my sister Sofía, Nelly, Celina, Sol, Amaya and Luis, Belén and all my nephwes.

Abstract

This thesis studies how firms' investment and credit are affected by different financial imperfections related to firm and bank learning, relationship lending and financial wealth. After reviewing in chapter 2 the related literature, in chapter 3 I investigate the main determinants of different types of financial constraints, such as credit rationing and excessive cost of debt, by constructing new measures of these problems based on qualitative data. I then develop in chapter 4 a model of firm investment with financial constraints and Bayesian learning that provides a new framework to analyze the problem of asymmetric learning between a bank and a firm and its effect on a firm's investment decision. This model is used to investigate, theoretically and empirically, the relationship between firms' investment and internal funds in the presence of limited information, learning and bankruptcy costs, providing new arguments to support a u-shaped curve theory of investment and internal funds. Finally, in chapter 5 this model is used to analyze how relationship lending affects the evolution of interest rates during the life cycle of firms.

Resumen

Esta tesis estudia cómo la inversión y el crédito están afectados por diferentes imperfecciones financieras relacionadas con el aprendizaje, las relaciones de crédito y la riqueza financiera. Luego de revisar la literatura relacionada, en el Capítulo 3 se investiga los principales determinantes de distintas restricciones financieras relacionadas con el acceso y las condiciones del crédito, mediante la construcción de nuevos indicadores de estos problemas. Luego, en el Capítulo 4 se desarrolla un modelo de inversión con restricciones financieras y aprendizaje Bayesiano que provee un nuevo marco para analizar el problema del aprendizaje asimétrico entre un banco y una firma y su efecto en las decisiones de inversión de esta última. Dicho modelo es utilizado para investigar de forma teórica y empírica la relación entre la inversión y los recursos propios en la presencia de información asimétrica, aprendizaje y costes de quiebra, obteniendo nuevos argumentos para apoyar la teoría de una relación en forma de U entre la inversión y los recursos propios. Finalmente, en el Capítulo 5 se estudia como una relación de crédito afecta la evolución de los tipos de interés durante el ciclo de vida de las firmas.

Table of Contents

CHAPTER 1.

1. INTRODUCTION

In the last 30 years a large body of theoretical and empirical literature has emphasized the importance of financial factors for economic stability and growth. The imperfections of the financial system may act as a mechanism of amplification of real and nominal shocks by limiting the ability of firms to access external finance and thus distorting their investment decisions. These frictions at the firm level may also impact the real sector at the aggregate level by amplifying cyclical fluctuations in real and nominal variables, for example by facilitating the creation of asset price bubbles during expansions and the subsequent banking and financial crises during recessions. Finally, financial development has an impact on economic growth and development through the reduction of financial frictions, which are usually more acute in less developed economies.

Financial frictions that hinder firms' growth are more severe for small, young and innovative firms. For example, these firms appear to be the ones that suffer tougher restrictions when credit is scarce after a financial crisis or after any contraction of economic activity. These financing problems have important aggregate implications. In the EU SMEs represented 99.8 % of all EU-27 enterprises in the non-financial business economy in 2006, employing two thirds of the workforce (67.4 %) and generating 57.7 % of total value added. [1] The concern that Small and Medium Enterprises (SMEs) are the most affected by credit restrictions and the documented importance of SMEs for employment and innovation has generated a lot of policy-oriented research and has driven worldwide public policies aimed at protecting and promoting access to credit to SMEs.

In order to have successful policies that boost financial development and reduce the impact of capital market imperfections on economic activity we should be able to properly understand what are the most important determinants of those imperfections at firm-level. To be able to do this, we should be able to understand and measure the effect of financial constraints on firms' real investment decisions. We should also be able to understand the reasons why some firms, such as SMEs, are more affected than other firms.

However, the existing research about the real effects of financial constraints has only provided partial answers to these important questions. For instance, there is still no generally accepted measure of financial constraints at the firm-level or at the macro level. More specifically, a wide range of theories have been proposed to microfound the presence of financing constraints on firm investment. Most of these theories about the possible determinants of access to credit and its cost show contradictory predictions and the related empirical evidence are also mixed and inconclusive.

This thesis studies how firms' investment and credit are affected by different financial imperfections related to firm and bank learning, relationship lending and financial wealth. The main added value of the thesis is to develop a new investment model with financial

[1] Source: Eurostat

constraints and Bayesian learning that provides new predictions about the investment behavior of firms and the interest rates they pay when creditors have a different perception of risk due to asymmetric learning, or when credit is too expensive due to bankruptcy costs. Importantly, the two different empirical exercises confirm the new theoretical predictions of the model. The first empirical exercise confirms the existence of a u-shaped relationship between investment and internal fund and it supports the explanations provided by the model. The second empirical exercise confirms the theoretical predictions about the evolution of interest rates during the life cycle of firms due to the change in creditors' perception of risk after the start of a lending relationship. These results can help to rationalise different mixed evidence obtained by earlier empirical literature. Besides the results based on the theoretical model, the thesis also investigates the main determinants of different types of financial constraints related to the access and the conditions of credit by constructing new measures of these problems based on qualitative data; The results provide new insights about the relationship between financial constraints and different firm and market characteristics.

The structure of the thesis is the following. In chapter 2 I review the most fundamental literature about financial constraints and investment. In section 2.1 I review the most influential studies about the real effects of asymmetric information on firms' investment, emphasizing those that introduced the problems of asymmetric information and agency issues into the credit market and the relevance of the firm's financial structure. In section 2.2 I review the most relevant contributions on learning and investment, especially those related to the differences in the observed financial behavior between firms of different characteristics. In section 2.3 I review the literature related to the identification and estimation of financial constraints. In section 2.4 I focus on the appropriateness of investment cash flow sensitivity as a measure of the presence and intensity of financial constraints. In section 2.5 I review different definitions of the concept of relationship lending. In section 2.6 I review the most important contributions to the literature about the effects of relationship lending on interest rates and access to credit and about the benefits and costs of relationship lending for the firm. Finally, in section 2.7 I review the literature on the relationship between bank competition and costs and access to credit, and in section 2.8, the literature on other determinants of financial constraints.

In chapter 3 I investigate the main determinants of different types of financial constraints related to low access to credit (credit rationing) and to unfavourable conditions in the use of bank credit, such as high interest rates and short maturities of bank loans. I perform this empirical research using direct survey-based information. Previous research on the determinants of financial constraints has been limited by the lack of clear measures of financial constraints and as a consequence has been unable to identify to what extent different firms are affected by different types of constraints, such as credit rationing versus high interest rates. This empirical analysis is an important step to identify the real effect of financing constraints. While many studies have used investment cash flow sensitivities as a measure of the intensity of financial constraints, other authors such as Almeida and Campello (2007) have pointed out that the investment-cash flow sensitivity of firms facing credit rationing could be completely different than that of firms facing high costs of external funds.

To obtain a more precise measure of the financial constraints faced by the firms is also important for the research on other related topics. For instance, it may help to understand what the effect of an increase in banking concentration on the access and cost of credit is, and more importantly, if banking concentration worsens or eases firms' financial restrictions. To date, both theoretical and empirical research has obtained mixed results exploring the effects of banking competition on the costs and availability of credit.

Other studies have used qualitative data to measure financial constraints. For instance, some surveys directly ask managers about their problems in their access to finance. However, as explained by Campello (2010), survey-based analyses also have limitations that are important to consider. One concern is that it is difficult to completely rule out a ‘‘state of mind’’ story that could somehow affect some respondents and not others. Survey-based inferences may also be compromised if respondents misjudge the economic conditions of their firms and misunderstand the way credit markets work. Therefore, it is necessary to verify that firms which say they are constrained also report tangible financing difficulties.

Chapter 3 addresses these issues and other problems found in previous studies, by applying a multivariate technique to a database that provides direct firm information on all types of firms around the world, the World Business Environmental Survey (WBES) 2000. The solution proposed is to combine those variables based on survey responses that could have ambiguous interpretations with other variables based on "hard data" information also provided in the same database. I thus use a multivariate technique called Categorical Principal Component Analysis (CATPCA) to combine the original variables into new indicators that summarize the characteristics related to the types of constraints we want to measure. Using the new constructed indicators, I am able to study which firm's characteristics and which credit market characteristics determine that firms have low access to bank credit or that they face unfavorable conditions of bank loans, such as high interest rates or short-term restrictions, while controlling for the different problems observed in the original variables of the survey.

By estimating two separated CATPCA analyses, I obtain the following results: With the first CATPCA model, I identify separately firms facing two different problems that cannot be differentiated with the original variables related to the access to credit. On the one hand I identify firms not having access to bank credit at all (credit rationed). On the other hand, I identify firms that do have access to bank credit, but only to short-term credit.

Similarly, with the second CATPCA model I identify separately firms facing different problems that cannot be differentiated with the original variable related to high interest rates. On the one hand, I identify firms that do have access to bank credit but that think that such credit is too expensive. On the other hand, I identify firms that do not have access to bank credit and are therefore forced to pay high interest rates to money lenders or suppliers. Finally, I identify firms that do not demand any type of credit because they think that is too expensive.

The new indicators are the principal components that summarize the expected characteristics of each one of the cases described before. After obtaining these new indicators, I run separate regressions with each one of them as dependent variables. The first result of the regression analysis indicates that access to bank credit clearly increases with firms' age and size and that the use of expensive sources of credit decreases as the firms grow and mature. However, as firms gain access to more bank credit, there is an initial increase in the perception of facing unfavorable conditions in the terms of the loans, such as too high interest rates and short-term restrictions. This initial increase reach a maximum point, after which the restrictions related to those unfavorable conditions start decreasing.

Therefore the results indicate that there is a hump shaped relationship between the perception of high interest rates of bank credit and the firm's age. This result has never been reported before for a measure of financial constraints or a measure of the perception of firms about too high interest rates. A similar evolution of the actual costs of credit has been reported by authors researching the effects of bank switching costs, as in Kim, Kristiansen and Vale (2008). Additionally, a similar hump shaped relation is found between the indicator of scarcity of long-term financing and the age of the firm.

Another important result is that banking competition (less concentration) improves firm's access to credit but it does not seem to have a significant effect on the perception of a high cost of bank credit. Another novel result is that although banking concentration worsens the general access of firms to bank credit, there is some evidence that it may somehow increase the access in terms of short term credit. These results seem to give partial support to the structure-performance hypothesis of banking concentration, since I find a positive relation between the main indicator of access to credit and banking concentration, but no relation between the main indicator of the costs of bank credit and concentration.

Similarly, different results were found with respect to other variables such as bankruptcy regulation or the level of credit information sharing in the market. More specifically, I find that in markets with more credit registry coverage, where it is easier to access borrowers' credit and financial information, firms have better access to credit (less credit rationing). I also find that a regulatory environment that strongly protects creditors' rights after bankruptcy reduces the restrictions related to the conditions of bank credit.

Chapters 4 and 5 of the thesis focus on the interrelations between learning, relationship lending, financial wealth, the demand for external funds and financial constraints. The main motivation for focusing on these topics was to provide a theoretical framework to understand the role of information in the perception of risk and the determination of the cost of credit. It is well known that SMEs and young firms are riskier than large corporations, and that such higher risk is important to understand why SMEs are more financially constrained. But few studies have investigated on how the different perception of risk between the firms and their creditors, in the presence of limited information and learning, could affect relationship lending (for a review of this literature, see section 2.2). Understanding how creditors process information and use it to determine credit conditions is key to understand how financing frictions affect the real decisions of firms.

The justifications for why small and young firms face the worst conditions in their access to external funds are usually: their opacity, lack of financial wealth or guarantees, and the asymmetry of information with respect to possible creditors and/or investors. Under certain conditions, asymmetric information or other financial frictions may force younger firms to pay higher interest rates for external funds or to have a limited access to them.

But paying higher interest rates it is not necessarily a sign of imperfect financial markets. On the one hand, a young or new firm may have very little information about its potential profitability to be able to estimate, in a precise manner, the chances of repaying a loan or completing an investment project successfully. A firm may only be able to do so after a long learning process. If a new firm itself has little information about its chances of repaying a loan, any possible creditor must also have scarce information to evaluate the default risk appropriately. In such case, it is reasonable to expect that a new firm pays a higher interest rate than an older one, but this has nothing to do with financial imperfections.

On the other hand, a bank may have little information about a firm's profitability not because the firm is young, but because it has had no previous contact with it and some important information may not be easy to transfer outside a relationship. In such case, the cost of external funds could be perceived by a firm to be higher than what it should be, if the bank were to have the entire firm's information (it could also be perceived as lower than what it should be).

In both cases, the bank may use the information acquired through its relationship with the firm to learn about the firm's prospects in a similar way that the firm learns about it. But in each case the effect of the bank's learning process may be different, and its effect on the conditions of access to external funds and on investment may be different as well. Thus, in order to understand the effect of the bank-firm relationship and financial constraints on firms' investment decisions we should separate the effect of asymmetry of information from the effect that the acquisition of more information have on uncertainty and the perception of risk. This is the main objective of chapters 4 and 5 of this thesis.

At this point it is important to clarify the concept of learning that is used throughout this thesis. Learning could be associated to a dynamic process in which learners improve their abilities through continued experience. A young, fast-growing firm can increase its profitability because managers improve their abilities and reduce their flaws by learning from experience. I do not consider this type of learning in the theoretical model developed in chapters 4 and 5. Rather, what I am implying by learning in this thesis is the process of accumulating information and increasing the knowledge about something. In the theoretical framework, firms and creditors update their beliefs about firms' intrinsic profitability as relationships mature, and banks gain access to longer firm histories.

In chapter 4 I develop a new framework for analyzing the possibility of asymmetric learning between a bank and a firm and its effect on firm's investment decisions. I apply this framework to investigate theoretically and empirically the relationship between firms' investment and internal funds in the presence of limited information, learning and bankruptcy costs. Even though many authors have shown that investment cash flow sensitivities are a very noisy measure of the intensity of financial constraints, these are still considered useful in evaluating the real effect of financing frictions, because of the lack of more precise measures. Recent empirical studies suggest that a firm's investment level is not a monotonically increasing function of internal funds, but instead that is a u-shaped function, because investment actually decreases with internal finance for the lowest levels of wealth.

To the best of my knowledge there is only one theoretical model that provides a possible explanation of why investment could be a u-shaped function of the level of internal funds, which is the model of Cleary, Povel and Raith (2007). These authors consider the model of a firm with idiosyncratic risk and with costly bankruptcy. However they derive the U-shape investment function under very restrictive assumptions, since their results only hold if the revenues shocks follow the uniform distribution.

In this chapter I show that the model with asymmetric learning is able to provide a more robust explanation of the U-shape in the investment-internal finance relation. More specifically, this chapter provides the following new results:

i) It shows that a U-shape investment function arises naturally in the presence of asymmetric learning of banks and firms about the firms' riskiness, even in the absence of bankruptcy costs.
ii) It extends Cleary *et al* (2007) model in order to clarify under what conditions bankruptcy costs are indeed sufficient for generating the U-shape function under more general distributional assumptions.
iii) It calibrates with realistic data both the model with asymmetric learning and the model with bankruptcy costs, to show that the U-shape function is more likely to arise if asymmetric learning is present.
iv) It estimates the U-shape function empirically with firm level data, to provide further evidence that asymmetric learning is the most likely explanation of it.

The main differences with Cleary *et al* (2007) are as follows: first, I assume that the bankruptcy cost is proportional to the value of investment and that, at the same time, firm revenues follow a log-normal distribution. Both assumptions are at the same time more realistic than Clearly *et al* (2007) assumptions and are sufficient to imply that bankruptcy costs still generate a U-shaped investment function[2].

Second, in this model I also consider a different type of capital market imperfection which is totally independent of the effect of the bankruptcy costs. I consider the possibility that creditors and firms do not share the same set of information for predicting the firm's probability of default. Therefore both firms and banks have to learn about the firm's profitability and about the probability of debt repayment in the presence of financial imperfections. More specifically, in the model firms and banks use historical information about a firm to learn and predict its future profits and default probability.

This chapter solves this model and shows that investment is non-monotonically related to internal funds in a u-shaped fashion. However, it is quite important to clarify that not all assumptions are needed jointly to generate the u-shape relationship. So by adding all these ingredients I define two possible explanations for the u-shape: the asymmetric learning and the non-monotonicity of the density function with respect to internal funds and its interaction with the expected bankruptcy costs.

The rationale for the u-shaped relationship in the case that the bank has a shorter history of information than the firm is the following: Any disagreement between a bank and a firm matters most when repayment or default are equally likely, and it matters less when any of

---

[2] Furthermore, the model does not require the unrealistic feature that when internal funds are negative debt could become cheaper the higher leverage is. Conversely, in my model I obtain a u-shaped investment curve in a setting in which, at the optimum, the firm cannot decrease its marginal costs by investing more, but on the contrary, if it invests more than its optimal level it would have to pay a higher interest rate because the probability of default would be higher.

them are almost certain. Intuitively, if the firm itself is not sure if it will default or if it will be able to repay its debt because it predicts that both situations are equally likely, any difference in the creditor's information could make a large difference in its prediction about the likelihood of those events, compared to the firm's beliefs. On the other hand, if either default or repayment is almost certain, the predictions from both sides are going to be quite similar independently of the set of information that they are using. Therefore, the effect of a disagreement between a creditor and a firm caused by the creditor's higher uncertainty depends on the level of internal funds: it is more important for intermediate levels of risk (intermediate levels of internal funds) and it is less important for high and low levels of risk (low and high levels of internal funds).

The rationale for the second possible explanation for the u-shaped relationship, the positive bankruptcy costs, is similar to the previous one, although totally independent of whether we assume asymmetry or symmetry of information: The effect of a marginal decrease in internal funds on the predicted probability of default is higher when repayment or default are similarly likely, and lower when any of them are almost certain. Therefore, the effect of a marginal change in internal funds and in firm's leverage depends on the level of internal funds: it is more important for intermediate levels of risk (intermediate levels of internal funds) and it is less important for high and low levels of risk (low and high levels of internal funds).

To test the theoretical predictions of the model, I first simulate an artificial sample of firms based on the calibration of the theoretical model with realistic parameters, in order to compare which one of the two alternative explanations for the u-shape investment curve is more likely to drive the empirical evidence. The results of the simulation imply that the asymmetric learning story is the one that more clearly could explain the existence of a u-shaped investment curve.

Finally, in the third part of the chapter I estimate a regression model of fixed capital investment using an unbalanced panel containing 257,566 firm-year observations of SABI data (SABI accounts for Iberia Balance Sheet System Analysis). Consistent with the theoretical predictions, all the empirical exercises support that investment is decreasing for the lowest values of internal funds and increasing for the highest levels.

The importance of the chapter results is mainly theoretical, but also empirical. The chapter provides new arguments to support the u-shaped curve theory of investment and internal funds and it does it in a more intuitive and general way than in previous studies. Additionally, it contributes to the general theory of investment and financial constraints by providing a completely new framework for analyzing the possibility of asymmetric learning and the estimation of probabilities with different information sets.

According to the predictions of the theoretical model, the empirical evidence confirms that young firms, which are the ones more likely to face stronger asymmetries of information and higher uncertainty in their distribution of profits, show a clearer u-shaped relation between investment and internal funds. Therefore, both the simulation results and the empirical evidence point in the direction of the asymmetric learning explanation. Importantly, I find that not all the firms with negative values of internal funds show a negative relation between investment and the different measures of internal funds.

In chapter 5 I investigate the effect that the length of the relationship between a bank and a firm has on the cost of credit and on the firm's demand for external funds and investment. Whether longer bank-firm relationships are beneficial or harmful for firms is an unresolved issue in the existing literature. Although some results on financial intermediation provide support for a positive value of relationship lending for small business finance, the "dark side" of relationship lending is that lenders endogenously gain an information monopoly and are able to extract additional revenue from borrowers in the form of higher interest rates in the later stages of the relationships. Besides the effect that such a monopoly of information could have on interest rates, there is evidently a facet of relationship lending that has not been explored enough in literature, which is the effect that gathering firm's information (learning) could have on the bank's estimation of the riskiness of the firm, and the effect that this learning process has on both the bank's and the firm's relative perception of risk.

Using the model of firm investment with financial constraints and Bayesian learning which I developed in chapter 4, I find that the process of banks' learning about firms, represented by the length of the relationship with the firm, may actually increase rather than decrease the interest rate paid by the firm when both the bank and the firm need to learn about the firm's quality. In this model, firms and banks need to use their history of information about the firm to learn and predict its future profits and default probability.

The model predicts that the total effect of a longer relationship on interest rates is the sum of two different effects that could have different signs: a reduction in uncertainty and a reduction in the disagreements between the firm and the bank when the bank has less information than the firm about the firm's intrinsic profitability. The reduction in uncertainty always generates a reduction in the interest rate via a reduction in the variance of prediction and the estimated probability of default. The reduction in the asymmetry of information and the possible disagreements between the firm and the bank has an effect on the loan's interest rate via a change in the firm's demand for capital and more specifically, on the demand for external funds and leverage. The latter effect is stronger for firms with relatively low levels of risk and is smaller for firms with high levels of risk.

The evolution of interest rates through the course of a bank-firm relationship depends on which effect dominates. When the effect on the demand for capital is low (for the high risk firms), we should observe that interest rates decrease or do not change as the length of the relationship increases. On the contrary, when the reduction in the demand for capital is high enough (for low and intermediate risk firms) we should observe that interest rates are initially very low (given the highly constrained demand of capital) and that they increase as asymmetry disappears and the firms demand more external funds at higher interest rates.

It is important to notice that this result is novel in the sense that an increase in interest rates throughout the course of a relationship does not come from possible lock-in problems that are claimed to generate monopolistic power for the bank through the acquisition of information over the course of a relationship.

I test these implications using the National Survey of Small Business Finance (NSSBF 2003) from US and the empirical evidence confirms the predictions of the model. First, I find that the estimated coefficients are consistent with a hump shape evolution of interest rates through the course of a relationship. However, if I restrict the sample to those firms with higher levels of leverage (higher risk) the estimated coefficients indicate a decreasing evolution of interest rates. Similarly, if I restrict the sample to those firms with medium and low levels of leverage, the estimated coefficients again support a hump shape evolution. Furthermore, the results confirm that such hump shape evolution is more likely to occur with asymmetry of information and that on the contrary, interest rates decrease with the length of relationship when there is symmetry of information.

There are two different empirical results that support the theoretical prediction of this chapter over those based on the lock-in effect or hold-up problems. First of all, models based on the lock-in problem do not predict a different evolution of interest rates depending on whether the length of relationship is equal or smaller than the age of the firm, meanwhile the model presented here predicts that interest rates decrease with longer relationships when the length is equal to the age of the firm and a hump-shape evolution when they are different; something that is confirmed with the data. Secondly, in a lock-in problem model the evolution of interest rates should be the same independently of the level of risk of the firm (more or less leverage or internal funds), meanwhile the empirical evidence presented herein shows that the effect is indeed different.

The results of the chapter are also important because they could help clarifying why the empirical results of the different studies that have studied the relationship between interest rates on commercial loans and the length of a relationship using similar US data is mixed and inconclusive. More importantly, it also adds the novel result that the final effect of the length of relationship depends on the degree of asymmetric information and on the degree of a firm's riskiness. Therefore, the main added value of the empirical analysis is to show that the effect of the length of relationship on loan characteristics can only be properly understood by considering the role of internal funds and the perception of risk on the demand for debt, and the role of the length relative to the age of the firm.

CHAPTER 2

2. LITERATURE REVIEW

2.1 Financial Constraints and Investment

In an extensive review of the theoretical and empirical literature, Levine (2005) concludes that a better functioning of financial systems and financial intermediaries mitigate the financial restrictions that hinder firms' growth and industrial expansion. In essence, the reduction of financial restrictions is the main mechanism through which financial development is relevant to economic growth and development.

In a perfect world with frictionless capital markets in which the Modigliani and Miller (1958) theorem applies, firms will choose their levels of investment in such a way that the marginal product of capital is equated across every project in the economy. In the real world, there are a variety of factors that prevent firms from doing so. This can have important effects on aggregated investment, and on the general level of activity. Taxes and transactions costs are examples of such frictions. But the most important sources of inefficiency of corporate investment are those that arise from informational asymmetries and agency problems.

Following the model classification by Stein (2003), the first broad class of models in the literature about asymmetric information and investment are those that unambiguously predict underinvestment relative to a first-best benchmark. In these models when firms have access to unlimited funds, investment converges to an efficient level. However, when firms do not have enough resources, investment becomes smaller than the first-best efficient level, because of the frictions associated with raising external finance. This group of models can be further sub-classified into models of high costs of equity finance, models of high costs of debt finance, models of credit rationing, models of debt overhang and models of incomplete contracts. The main models of high costs of equity finance, debt overhang and those that describe the agency problems between managers and stockholders are reviewed in Stein (2003).

The first models of asymmetric information and credit rationing (Jaffee and Russell (1976), Stiglitz and Weiss (1981, 1983) show that adverse selection and moral hazard can lead to credit rationing, making firms simply unable to obtain all the debt financing they would like at the equilibrium market interest rate. The key insight in this line of work is that high-quality borrowers (those with a low probability of default) may have difficulty at conveying information about their credit quality to lenders. As long as that the debt involved has some default risk, firms will be more likely to borrow if their private information suggests that they are more likely to default than what creditors may expect, at any given interest rate. In certain cases, some (high-quality) borrowers are unable to obtain funds at any interest rate. Credit rationing can occur because of a moral hazard problem, because those firms who borrow have an increased incentive to take the higher risks after issuing new debt, increasing the risks of default ex post, which can weaken firms' access to external funds, as in Gertler (1992), Boyd and Prescott (1986) or Diamond (1984).

One group of models that is closely related to this thesis is the one related to high costs of debt finance and costly state verification, for instance, Townsend (1979) and Gale and Hellwig (1985). They assume that outside investors can only verify a firm's cash-flows by paying some fixed auditing cost. As long as the firm repays its debt, there is no audit, and the manager gets to keep the rest of the firm's cash-flow. However, if the manager fails to make the debt payment, the lender audits, and keeps everything he finds; this can be interpreted as costly bankruptcy. These models usually imply that there is a degree of underinvestment that is directly related to the auditing/bankruptcy cost and the likelihood of the auditing/bankruptcy cost being incurred. In particular, the less wealth the firm is able to put up and hence the more he must borrow, the greater the deadweight costs of external finance and the smaller the investment.

Another important group of models is composed by those related to incomplete contracts and imperfect enforceability. Similar to the costly-state-verification models, incomplete contracts models also have the feature that there is underinvestment ex ante, with this problem being a decreasing function of firms' wealth. Debt is often seen as an incentive scheme that rewards firms with continued control if they repay their debt on time and punish firms with loss of control otherwise. In a multi-period framework, this type of incentive scheme enables outside lenders to extract payments from firms even in the extreme case where cash-flows are completely unverifiable. Well-known papers in this vein include Aghion and Bolton (1992), Bolton and Scharfstein (1990) and Hart and Moore (1994, 1998).

Albuquerque and Hopenhayn (2004) add the features that leverage and revenues are state contingent and allow for liquidation of the firm, which gives rise to non feasibility of contracts, credit rationing and underinvestment, and inefficient termination of projects.

2.2 Learning and Investment

The literature about learning and investment has focused on explaining the different observed financial behavior between firms of different characteristics (like firm's age) that could be related to learning, such as the relation between size and age and the reported higher sensitivity of investment to internal funds of smaller firms.

Some empirical papers in the industrial organization literature show that younger and smaller firms grow faster, and their growth rates are more variable (Evans (1987), Hall (1987)). The higher variability of the growth rates of younger firms gives support to the idea that these firms learn about uncertain growth prospects through time. Some theories of firm dynamics and firm life cycle that do not include financial frictions are based in the idea that young firms need to learn about their productive capabilities and this learning process drives the differences in financial behavior between young and mature firms (Jovanovic (1982), Weinberg (1994), Li and Weinberg (2003)).

Some authors suggest that the process of firm learning explains in a better way the observed differences in financial behavior and therefore, weakens the role of financial constraints. For instance Li and Weinberg (2003) propose that young firms grow faster than mature firms and have lower survival rates not because they face higher financial constraints but because young firms still need to learn about their own profitability.

As described by Weinberg (2000), we could find two main theoretical perspectives on firm size and finance that intends to explain the evidence provided by the many empirical studies that show how small and large firms differ. The focus of the first is on imperfections in financial markets. The second focuses on the causes of variations in firm sizes in a dynamic, competitive economy.

Weinberg argues that in a full information, Modigliani-Miller world, a firm would be indifferent between the use of internal and external funds. But if in a full information environment there were small transactions cost associated with raising external funds, firms would have sufficient reason to prefer internal funds. That is, rather than paying dividends and raising funds externally, a young firm with good growth prospects will retain earnings to fund its likely investment needs. Hence, problems of asymmetric information are sufficient but not necessary for a preference for internal funding.

For Weinberg, a learning model is consistent with the observations on investment behavior and the use of internal funds. Small firms are more likely to be young firms and engaged in learning. For these firms, the presence of favorable investment opportunities is correlated with the presence of ample internal funds, generated from current and recent favorable performance. Larger firms are more likely to be mature. For these firms, investment opportunities are less tied to firm-specific learning from experience. They are correspondingly more likely to have opportunities arise in times of low internal resources, requiring them to look for external sources for funds.

Alti (2003) develops a model in which firms learn about project quality in a similar way as the learning process in Jovanovic (1982). In Jovanovic's model, a firm learns its efficiency from cost realizations. In his model, firms do not have capital stocks and they can freely set their output. Jovanovic's focus is on the endogenous selection that occurs due to low quality firms exiting the industry. By analyzing the case of high-growth firms learn their projects' qualities through time (i.e. where information content of cash flow follows a non-stationary process), Alti (2003) shows that the excess sensitivity of investment to cash flow may obtain even in a frictionless environment.

In summary, for authors such as Weinberg and Alti (2003) a theoretical perspective based on asymmetric information that produces financial constraints is capable of explaining observed deviations from the type of behavior predicted by the frictionless framework of Modigliani and Miller. However, they claim that this perspective cannot fully explain how those deviations tend to be more apparent for smaller than for larger firms. For Weinberg, the explanation of why the asymmetric information problems weigh more heavily on some firms can be found in a life cycle perspective. As firms age and grow, they acquire a public reputation that can partially undo the constraints imposed by informational frictions. For him, the life cycle perspective is capable of explaining a great deal of the observed behavior by itself.

Clearly the two theoretical approaches are not mutually exclusive. Firms that are young and still accumulating knowledge about themselves are likely to be firms about which insiders are better informed than outsiders.

In this thesis I study how we can integrate financial frictions such as asymmetry of information, transaction/bankruptcy costs and the learning process of firms and creditors to have a broader picture of firms’ financial behavior.

2.3 Measures of Financial Constraints and Investment Cash-Flow Sensitivity

Financial Constraints related to asymmetric information problems are inherently unobservable or in the case of other frictions such as bankruptcy or monitoring costs are very difficult to quantify. Researchers have followed very different ways to measure financial constraints. The most common methods are based on inferring the presence of financial constraints from certain observable characteristics of firms, such as the sensitivity of investment to internal funds or the use of indexes of financial characteristics associated to financial weakness, as in Cleary (1999) or Whited and Wu (2006). The third most used methodology is the use of surveys, directly asking firms about their problems in the access and use of external finance, as in Beck et al (2004), Beck et al (2006) and Campello et al (2010).

A large part of the research about financial constraints and investment has focused on the investment-cash flow sensitivity debate started by Fazzari, Hubbard and Petersen (1988) that aim to measure financial constraints based on the idea that when financial imperfections are present, investment must be sensitive to changes in internal funds. This line of research started by recognizing that imperfections in capital markets may lead to a wedge between internal and external cost of funds as a consequence of informational asymmetries, holding investment opportunities constant.

With asymmetric information about the risk and quality of firms' investment projects, adverse selection leads to a gap between the cost of external financing in a uninformed capital market and internally generated fund (e.g., Greenwald, Stiglitz and Weiss (1984), Myers and Majluf (1984)). Alternatively, in the presence of incentive problems and costly monitoring of management behavior, external investors require a higher return to compensate for the monitoring cost and the potential moral hazard associated with managers' control over the allocation of investment funds (e.g., Bernanke and Gertler (1989,1990)).

Fazzari et al (1988) and all the studies that followed their methodology derive an empirical specification from the firm's investment equation that describes the firm's optimal investment condition. One model is the q-model of investment, pioneered by Tobin (1969) and extended to models of investment by Hayashi (1982). Financial frictions are introduced to the model by adding internal funds variables such as cash flow.

Fazzari et al (1988) find that investment is positively related to firms' cash flows and that the coefficient is larger for firms with low dividend payouts relative to those with high payouts. Subsequent studies support their central finding, if they split samples according to certain a priori measures of financial constraint.

Kaplan and Zingales (1997) classify firms into categories of ''not financially constrained'' to ''financially constrained'' based upon statements contained in annual reports. They classify firms as being severely financially constrained if these companies are in violation of debt covenants, have been cut out of their usual source of credit, are renegotiating debt payments, or declare that they are forced to reduce investments because of liquidity problems. They claim that oppositely to Fazzari et al findings, financially constrained firms show lower investment cash-flow sensitivity than unconstrained ones. Therefore, their results contradict

Fazzari et al.'s findings that more financially constrained firms have higher investment-cash flow sensitivities.

Cleary (1999) finds similar evidence in a large sample of manufacturing firms, suggesting that Kaplan and Zingales' results might not be driven by their small, homogenous sample. Furthermore, the simulations in Gomes (2001), Alti (2003) and the empirical work in Erickson and Whited (2000), Bond and Cummins (2001), and Cooper and Ejarque (2003) have demonstrated that observed cash-flow sensitivity is likely an artifact of measurement error in the Tobin's q, which is the most common proxy for investment opportunities. In sum, these papers suggest that cash flow is correlated with investment, but that this correlation is not necessarily an indication of finance constraints.

An alternative approach to simply introducing a variable for internal funds into a standard investment regression is to derive an empirical specification from the firm's investment Euler equation describing the firm's optimal investment pattern. The Euler model of investment has been applied and further developed by Abel and Blanchard (1986), Bond and Meghir (1994), and Gilchrist and Himmelberg (1995, 1998), among others. A third approach, introduced by Demirguç-Kunt and Maksimovic (1998), estimates a financial planning model to obtain the maximum growth rate firms can attain without access to external finance. By comparing these growth rates with the actual growth rates of firms, they are able to infer the degree to which firms are financially constrained.

A more recent line of research explains that the different results found by the seminal papers of Fazzari, Hubard and Petersen (1988) versus Kaplan and Zingales (1997) and Moyen (1999) can be attributed to a non-monotonic relationship between internal funds and investment. Allayannis and Mozumdar (2004) find that the mixed empirical results can be attributed to the negative cash flow observations. They suggest that firms with very low levels of internal funds (negative) are able to make only the absolute essential investment, and therefore, investment is less sensitive for firms with very low or negative cash flows.

Lyandres (2007) also finds that by altering the optimal investment timing, costly financing affects current investment and the sensitivity of investment to internal cash flow, making the relation between the cost of external funds and investment–cash flow sensitivity, non-monotonic.

Almeida and Campello (2002) find that if we considered being financially constrained as being credit rationed, the predicted relation between investment and cash flow is different from the one that was usually assumed in most of the investment empirical research. They showed that the less credit-rationed firms might have higher investment-cash flow sensitivities than the more constrained ones.

Finally, in a closely related paper to the one proposed in chapter 4, Cleary, Povel and Raith (2007) find that a firm's investment level is a u-shaped function of its internal funds. They show that for sufficiently low levels of internal funds, a further decrease in internal funds leads to an increase in the firm's investment. Their theoretical model shows that, in particular, investment is a decreasing function of internal funds for an important fraction of firms with negative levels of cash flow.

After the papers of Kaplan and Zingales and the doubts about the usefulness of cash flow sensitivities, one common approach to measuring financial constraints is to construct new variables (indexes) composed of different financial variables that may indicate the presence of financial constraints. The earliest example is that of Cleary (1999), who determines financial status using multivariate discriminant analysis, similar to Altman's Z score for predicting bankruptcy. According to Cleary, his classification scheme effectively captures the desired cross-sectional properties of different firms.

Whited and Wu (2006) estimate an index of firms' external finance constraints via GMM estimation of an investment Euler equation. Unlike the KZ index of Cleary (1999), their index is consistent with firm characteristics typically associated with limited access to external financial markets. They base their index of financial constraints on a standard inter-temporal investment model augmented to account for financial frictions.

Finally some studies have tried to avoid having to infer financing constraints from financial statements of firms as in Fazzari et al. (1988) and Kaplan and Zingales (1997) by using direct survey data on firms' perceptions about their access to external finance and the problems related to it, such as high interest rates, low access to bank credit or too high collateral requirements, as in Beck, Demirguç-Kunt, Laeven and Maksimovic (2004, 2006).

The most recent example of this line of research is that of Campello, Graham and Harvey (2010), who survey 1,050 Chief Financial Officers (CFOs) in the U.S., Europe, and Asia to directly assess whether their firms are credit constrained during the global financial crisis of 2008. They study whether corporate spending plans differ conditional on this survey- based measure of financial constraint. Their evidence indicates that constrained firms planned deeper cuts in tech spending, employment, and capital spending. Constrained firms also burned through more cash, drew more heavily on lines of credit for fear banks would restrict access in the future, and sold more assets to fund their operations. They also find that the inability to borrow externally caused many firms to bypass attractive investment opportunities. More than half of the respondents said they canceled or postponed their planned investments.

However, as described by Campello et al (2010), survey-based analyses have limitations that are important to consider. One concern is whether some respondents (CFOs in their case) simply ''perceive'' credit to be scarce and invest less anticipating a demand contraction in the crisis. Given that it is difficult to ultimately rule out a ''state of mind'' story that could somehow affect some CFOs and not others, it is necessary to verify that firms which say they are constrained also report tangible financing difficulties. A related concern is that respondents by themselves may not be able to separate economic from financial effects when responding to a survey. Survey-based inferences will be compromised if respondents misjudge the economic conditions of their firms and misunderstand the way credit markets respond, a possibility that cannot be ruled out. Finally, another concern related to survey-based analysis is whether uncontrolled firm heterogeneity could confound the inferences.

2.4 Relationship Lending

One of the most important ways in which financial intermediaries help reducing financial restrictions and add value to the firms is the production and collection of firms' information. And the easiest way in which financial intermediaries can produce and accumulate relevant information is through the continuous contact and the provision of products and services to the firm, a concept that has been called relationship banking or relationship lending.

Relationship banking (lending) represents an investment process from the banks and possibly from the firms that bring in benefits for both of them. Banks collect firms' information because information about borrowers is vitally important to the lending process. At the same time, the information collected by the banks throughout the relationship could be very valuable for the firm because it helps reducing possible information asymmetries and can be very helpful in the credit decision process (see e.g., Diamond, 1984, 1991).

As explained by Freixas (2005), the term relationship banking is not rigorously defined in the literature. Freixas uses this term as referring to the investment in providing financial services that allow to repeatedly dealing 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, although relationship banking may include the provision of other financial services rather than just lending.

Ongena and Smith (2000) define a bank relationship to be the connection between a bank and a customer that goes beyond the execution of simple, anonymous, financial transactions". Boot (2000) also provides a definition of relationship banking as "the provision of financial services by a financial intermediary that: (i) invests in obtaining customer-specific information, often proprietary in nature; and (ii) evaluates the profitability of these investments through multiple interactions with the same customer over time and/or across products". More recently, Elsas (2005) defines relationship lending as a "long-term implicit contract between a bank and its debtor".

Lenders produce or gather information about the borrower beyond the information that is readily available. This information is obtained ex ante in the screening process and during the relationship, with continued monitoring, and provision of multiple financial services. This information can be costly to acquire but sometimes the information is generated as a free by-product of the relationship. An additional motivation for the agents to invest in the search for information is being able to reuse the information in the future. The most recent contributions give special emphasis to banks gathering "soft information", i.e. generally non-quantifiable information obtained through interactions with the firm, its owner, suppliers, customers and the community.

2.5 Effects of Relationship Lending on Interest Rates and Access to Credit

It is reasonable to expect that the longer a relationship exists between a borrower and a lender the greater the information flow between the two parties and the more important the relationship becomes. However, whether a longer relationship is beneficial or harmful for the firms is, however, an unresolved issue in the existing literature.

Relationship lending adds value to the firm because facilitates the information exchange between the borrower and the lender. Lenders invest in generating information from their client firms and borrowers are more inclined to disclose information because of the preservation of certain confidentiality (Yosha 1995). The lower informational asymmetries make it possible to overcome problems of moral hazard and adverse selection otherwise inherent in credit markets. For instance, they ameliorate the project-choice moral hazard (Diamond 1991) and solve agency problems of managerial behavior (Weinstein and Yafeh 1998).

Another benefit of relationship lending that has been highlighted in the literature is that repeated lending from a bank provides credible certification of payment ability. This permits borrowers to build a reputation that would allow eventual borrowing through public markets (Fama 1985, Diamond 1991).

Relationship lending allows for loan contracts that are welfare enhancing which otherwise could not be contractible. Boot (2000) argues that relationship lending allows for implicit long-term contracts, more flexibility in renegotiation and some discretion in order to make use of soft information disclosed during the relationship.

The line of research that is the most related to the work presented in this thesis is the one that has focused on the evolution of interest rates or the access to credit over the time or the duration of a relationship. This literature has specialized in modeling the characteristics of loan contracts and how these characteristics may vary with the length, the scope, or the number of relationships, and with the credit markets conditions, such as the competitiveness of the markets or the size of the banks.

With respect to the effect of the length of relationship on loans' interest rate, Boot and Thakor (1994) show that even without learning or risk aversion, loan rates decline as a relationship matures. The intuition is that the long-term contracting permitted by a durable relationship enables the bank to efficiently tax and subsidize the borrower through time to reduce the use of (costly) collateral. Prior to the first project success, the borrower must accept a secured loan with an above spot market borrowing cost. But after the period in which the borrower encounters this first success, he is awarded an unsecured loan with a below spot market interest rate in every subsequent period perpetually.

On the other hand, Greenbaum et al. (1989), Sharpe (1990), Rajan (1992) and Wilson (1993) predict that loan rates increase with the duration of a bank–firm relationship. Bank relationships produce an asymmetric evolution of the information between the relationship bank, who acquires private information on the borrower, and the rest of financial intermediaries outside the relationship. This results in an informational monopoly (“hold-up") of the former. Sharpe (1990) and Rajan (1992) argue that informed banks endogenously gain bargaining power and are able to extract monopoly rents from borrowers. They argue that the bank’s improved knowledge locks the borrower into the relationship, enabling the bank to charge above-cost interest rates as the relationship continues.

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 as in Klemperer (1987). Switching costs models also predict that interest rates may increase with the duration of a relationship due to a lock in effect or an increase in monitoring costs. For instance, Kim, Kristiansen and Vale (2008) find that Banks’ interest rate markups are predicted to follow a life-cycle pattern over the borrowing firms’ age due to endogenous bank monitoring by competing banks. In their model, interest rates first increase with the age of the firm and decrease afterwards.

In a related venue of research, Petersen and Rajan (1994,1995) investigated the effect of relationship lending on the availability of credit and on the cost of credit. They do not find a significant effect of the length of relationship on the interest rate, but they found a different marginal effect of the age of the firm on the interest rates depending on the level of competition in the credit market the firm is located.

They argue that the level of competition on the credit markets may affect the incentives of banks to lend to the firms. In concentrated markets a bank has an incentive to offer cheap credit at the beginning of a relationship given that because of the lack of competition the firm will be lock-in into a relationship and could be forced to pay higher interest rates after a relationship has been established. Therefore, relationship lending increase credit availability, in particular to the youngest and informationally opaque borrowers, which may have projects that generate few rents in the first period but may be profitable from a long term perspective (Petersen and Rajan 1995). Even more, relationships permit smoothing the loan interest rate over the duration of the relationship (Petersen and Rajan 1995) and over the interest rate cycle (e.g. Berlin and Mester 1998, Ferri and Messori 2000).

The evidence from the empirical literature about the relation between interest rates on commercial loans and the length of a relationship is also mixed and inconclusive. Following two surveys of the literature from Ongena and Smith (2000), and Elyasiani and Goldberg (2004), only Berger and Udell (1995) have found a significant and negative effect of the length of a relationship on the interest rates paid using data from the NSSBF. They find a significant negative effect of the length of relationship on interest rates, but only using lines of credit and only on those loans with floating interest rates. They also find a significant negative effect of the length of relationship on the probability of pledging collateral.

On the other hand, Degryse and van Cayseele (2000) find a positive effect of length on interest rates using a large survey of Belgian firms. Blackwell and Winters (1997) do not find a significant effect of the length on interest rates. They, however, find a negative effect of the length on monitoring frequency. Elsas and Krahnen (1998) and Harhoff and Körting

(1998) neither find a significant effect of length on cost of credit. Finally, Angelini, Di Salvo and Ferri (1998) find a positive effect of length on interest rates using Italian survey data.

Besides the literature that has explicitly explored the effect of the length of relationship on interest rates or access to credit, several papers that have studied other related topics such as race and gender discrimination in the small credit market, or the distance from a bank, have not reported a significant relationship between the length of the relationship and the interest rate paid on the more recent loan using different versions of the NSSBF. (For instance, Blanchflower, Levine and Zimmerman (2003), Cavalluzzo and Wolken, (2002), Petersen and Rajan (2002)). Some of these papers do not even report a significant effect of firms' age on interest rates.

2.6 Bank Competition and Costs and Access to Credit

In the economic literature we can find two different hypotheses about the effect of bank competition on the firms' conditions in their access to finance. The structure-performance hypothesis predicts a negative relation between bank concentration and access to credit, while the information-based hypothesis predicts a positive or nonlinear relation.

On the one hand, the structure-performance hypothesis suggests that any deviation from perfect competition results in less access by borrowers to loans at a higher cost. Using an endogenous growth model, Pagano (1993) interprets the spread between lending and deposit rates as reflecting "the X-inefficiency of the intermediaries and their market power". Concentration in the bank market reflects less competition and more market power for the main banks. This market power and market inefficiencies are translated into a wider spread between lending and deposit rates. Guzman (2000) shows that a banking monopoly is more likely to result in credit rationing than a competitive banking market and leads to a lower capital accumulation rate. Boot and Thakor (2000) suggests that competitive markets may be beneficial to relationship building.

On the other hand, the so called "information hypothesis" claims that in a dynamic framework, certain informational asymmetries can create incentives for banks with more market power to lend to more informational opaque borrowers, since in a less competitive environment banks can establish long-term relationships with young borrowers and share in future surpluses. For instance, Petersen and Rajan (1995) showed that creditors are more likely to finance credit constrained firms when credit markets are more concentrated because it is easier for these creditors to internalize the benefits of assisting the firms. Banks with market power have more incentives to establish long-term relationships because they can share in future surpluses. Similarly, Marquez (2002) shows that borrower-specific information becomes more disperse in more competitive banking markets, resulting in less efficient borrower screening and most likely in higher interest rates.

There are also models that imply that there could be offsetting effects of bank concentration. For instance, Dinç (2000) shows that there is an inverted U-shaped relation between the amount of relationship lending and the number of banks, with an intermediate number of banks able to sustain the maximum amount of relationship lending. Similarly, Cetorelli and Peretto(2000) show that on one hand bank concentration reduces the total amount of loanable funds, but on the other hand it increases the incentives to screen borrowers and thus the efficiency of lending. The optimal banking market structure is thus an oligopoly rather than a monopoly or perfect competition.

In line with the different implications of theoretical models, empirical studies offer mixed results. Some studies find that higher concentration is associated with higher credit availability consistent with the information hypothesis that less competitive banks have more incentive to invest in soft information (e.g., Petersen and Rajan, 1995 and Zarutskie, 2004).

Other empirical studies, however, find support for the market power hypothesis that credit rationing is higher in more concentrated bank markets and that competition enhances access to credit (e.g., Jayaratne and Wolken, 1999; Boot and Thakor, 2000; Ongena and Smith, 2001, and Scott and Dunkelberg 2005, Elsas 2005). On the other hand, other papers have found evidence consistent with the information hypothesis. The methodologies and the data sets reflected in this literature vary considerably.

An interesting line of research is the one that 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 (1998) and Dell'Ariccia (2001) obtain instead ambiguous results regarding the effect of competition on relationship banking.

Most empirical studies of the effect of bank concentration on access to external finance and firm growth have focused on individual countries and mostly the U.S. Hannan and Berger (1991) finds strong evidence that concentration is associated with higher interest rates across U.S. banking markets. Similarly, Black and Strahan (2002) find evidence across U.S. states that higher concentration results in less new firm formation, especially in states and periods with regulated banking markets. DeYoung, Goldberg, and White (1999) find across local U.S. banking markets that concentration affects small business lending positively in urban markets and negatively in rural markets. Finally, using survey data from a panel of small U.S. firms, Scott and Dunkelberg (2001) find that a Herfindahl index of bank concentration is not robustly correlated with the availability and cost of credit, while a firm-based assessment of the competitive environment is.

Cetorelli and Gambera (2001) use industry-level data for 41 countries and find that while bank concentration imposes a deadweight loss on the overall economy, it fosters the growth of industries whose younger firms depend heavily on external finance. However, this positive effect is offset in banking systems that are heavily dominated by government-owned banks. Using a similar model, Cetorelli (2004) shows that financially dependent industries are more concentrated in countries with more concentrated banking systems.

One of the most closely related studies to the one presented in Chapter 3 is Beck et al (2004) that using data from the World Business Environmental Survey (WBES) assess the effect of banking market structure on the access of firms to bank finance They find that bank concentration increases obstacles to obtaining finance, but only in countries with low levels of economic and institutional development. They use a survey response indicator of financial constraints, which measures whether firms find that financing is (in general) an obstacle for the firm.

Overall, both theoretical and empirical contributions yield contradictory conclusions. Further, the relation might vary for firms of different sizes and across different institutional environments and ownership structures of the banking system.

The resolution of these conflicting views is not only interesting from the perspective of understanding the nature of relationship lending, it also interesting because the issue of the competitiveness of the global banking industry has become an central issue given the possibility that the global consolidation of the banking industry could produce a less competitive commercial loan market. Of particular concern is the prospect that consolidation could lead to a contraction in the number of banks that specialize in relationship lending, such as smaller community banks.

2.7 Other Determinants of Financial Constraints

The research about the determinants of financial constraints has been limited by the difficulty of measuring financial constraints and the debate about the usefulness of the cash flow sensitivity of investment. Although a generally accepted measure has not been developed, certain firms' characteristics have been a priori associated by the literature as related to higher informational asymmetries and thus to financial frictions. For instance, several empirical studies have grouped firms by dividend payouts (Fazzari et al., 1988), business-group affiliation (Hoshi et al., 1991), size and age (Devereux and Schiantarelli,1990), the presence of bond ratings (Whited, 1992), the degree of shareholder concentration, or the pattern of insider trading (Oliner and Rudebusch, 1992).

By using these a priori classifications and relying on investment cash flow sensitivities as appropriate measures there is a wide variety of studies that have found different firm and country determinants of financial constraints. Hoshi et al. (1991) find that investment is less sensitive to cash flows for firms that are members of a keiretsu and are presumed to be less financially constrained. Oliner and Rudebusch (1992) show that investment is more closely related to cash flow for firms that are young, whose stocks are traded over-the-counter, and that exhibit insider trading behavior consistent with privately held information. Demirguç-Kunt and Maksimovic (1998) find that financing constraints are lower in countries with more efficient legal systems. Love (2003) finds a strong negative relationship between the sensitivity of investment to the availability of internal funds and an indicator of financial market development, and concludes that financial development reduces the effect of financing constraints on investment. Laeven (2003) and Gelos and Werner (2002) find that financial liberalization relaxes financing constraints of firms, in particular for smaller firms.

Beck et al (2006) use survey-based indicators of financial constraints to assess the impact of different firm and country-market characteristics as determinants of financial obstacles. They also explore whether financial and economic development helps alleviate the financing obstacles of the firms that report to be most constrained. They find that older, larger, and foreign-owned firms report less financing obstacles. Their results also suggest that firms in countries with higher levels of financial intermediary development, stock market development, legal system efficiency and higher GDP per capita report, on average, lower financing obstacles, showing that the underlying institutions driving both financial and economic development seem to be the most important country characteristic explaining cross-country variation in firms' financing obstacles.

Winker (1999) uses a micro data from the German IFO institute to assess the relevance and impact of credit rationing at the firm level. The data used allows him constructing a variable closely correlated with the informational asymmetry between the firm and its possible creditors. The empirical results suggest that, ceteris paribus, older firms face a lower risk of being rationed on the credit market, whereas asymmetric information due to improving business expectations increases the risk of being rationed.

Related to the claim that relationship banking is valuable to the firm given that it reduces informational asymmetries, the empirical literature about relationship lending has highlighted three important determinants of relationship banking that could be also associated to the presence of financial constraints: firms' age, size and the type of business they perform. The age and size of a firm are 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, 2001).

Hovakimian (2009) finds that investment-cash flow sensitivity is non-monotonic with respect to financial constraints, cash flows, and growth opportunities. Specifically, firms with negative cash flow sensitivity have the lowest cash flows and highest growth opportunities, and appear the most financially constrained. Cash flow insensitive firms have the highest cash flows and lowest growth opportunities, and appear the least financially constrained. At least partially, negative cash flow sensitivity is driven by high investment and low cash flow levels at the inception of firms as public companies, which decrease and increase, respectively, with age.

CHAPTER 3

3. ON THE DETERMINANTS OF DIFFERENT TYPES OF FINANCIAL CONSTRAINTS: CREDIT RATIONING AND HIGH COSTS OF DEBT

ABSTRACT

This chapter investigates the main determinants of different types of financial constraints related to low access to credit (credit rationing) and to restrictions in the conditions of credit, such as high interest rates and short maturities. I first use a multivariate technique to construct appropriate measures of firms facing these types of constraints, based on survey responses of firms around the world (World Business Environmental Survey WBES). Using these new variables I can determine the main firm and market determinants of the presence and intensity of these different constraints. I find that the effect of firm's age and size on access to credit is different than the effect on constraints related to the costs and conditions of bank loans. Younger and smaller firms are constrained mainly in terms of not having enough access to credit (credit rationing), while constraints in terms of high interest rates and short maturities are more related to medium-age and large firms. Banking concentration reduces access to credit, but it does not have a significant effect in terms of increasing its cost. Similar different results were found with respect to other variables such as bankruptcy regulation or the level of credit information sharing in the market.

3.1 Introduction

As explained in the general review of the literature, the research about the determinants of financial constraints has been limited by the lack of a generally accepted empirical measure of financial constraints. A widely used measure is the investment sensitivity to cash flow. However, a great deal of the research has been around defining whether investment cash flow sensitivities could even be accepted or not as evidence of the presence of financial constraints, as it is reviewed in section 2.3. Given the doubts about the validity of this measure, the conclusions of the studies that have investigated the determinants of financial constraints using investment cash-flow sensitivities as an indicator of financial constraints have to be taken with caution.

Other few studies have used survey-based information about financial constraints to explore the determinants of firms' financing obstacles. Based on the same database that I use in this chapter (World Business Environmental Survey), Beck et al (2004) and Beck et al (2006) explore the main determinants of certain financing obstacles of firms around the world.

On the one hand, Beck et al (2004) investigate the effect of banking concentration on a variable that measures whether firms find that financing is (in general) an obstacle for the firm. However, the variable used as indicator does not allow exploring whether banking concentration and other market structure characteristics could have different effects on different types of constraints, such as access to credit versus high interest rates. Since different theories about banking concentration and relationship banking make conflicting predictions about the effects of concentration on access to credit and the cost of credit, it would be important to study the effect of concentration on each variable separately, to be able to test the implications of each different theory. The structure-performance hypothesis suggests that any deviation from perfect competition results in less access by borrowers to loans at a higher cost, as in Pagano (1993), Guzman (2000) and Boot and Thakor (2000). On the contrary, the so called "information hypothesis" claims that in a dynamic framework, certain informational asymmetries can create incentives for banks with more market power to lend to more informational opaque borrowers, since in a less competitive environment banks can establish long-term relationships with young borrowers and share in future surpluses (through higher interest rates), as in Petersen and Rajan (1995) and Marquez (2002).

On the other hand, Beck et al (2006) explore which firms' characteristics are related to different proxies of financial constraints, but they do not include country characteristics among the explanatory variables of those different types of constraints. More importantly, the way the questions related to high interest rates and access to credit were phrased in the WBES survey may limit their interpretation and the conclusions we can extract from them. First of all, one of the questions in the survey asks whether the access to long-term loans is an obstacle for the firm, but the question does not allow knowing whether a firm does not have access to bank credit at all or if it only has access to short term loans (cannot find long-term financing). Since the first problem could be more restrictive for a firm, it would be important to know the determinants of each problem. Additionally, the variable related to high interest rates has also different problems. The question only says whether high interest rates are an obstacle for the firm, but we do not observe the actual interest rate paid; neither can we observe the type of credit that the firm is actually complaining about. Therefore, if a firm complains that high interest rates are an obstacle we cannot know if it claims so because a bank is asking the firm very high interest rates or if a firm that does not have access to bank credit is forced to pay high interest rates to money lenders or suppliers, a problem that can generate erroneous conclusions.

For instance, among other results, Beck et al (2006) find that small and medium firms are the ones with stronger complaints about high interest rates and access to long term loans. Although it does not seem surprising at the first sight, the result that small firms are the ones with higher complaints about interest rates could be used as an argument to support a public policy of interest rates caps on formal credit from banks and other regulated financial institutions. However, interest rates caps on bank credit has been associated to a lower supply of formal credit to small firms and a larger use of informal credit from unregulated agents such as money lenders. Additionally, the use of certain expensive sources of funds such as supplier credit has been associated to the lack of access to bank credit. For instance, Petersen and Rajan (1995) utilize the use of supplier credit as a proxy for credit rationing. Therefore, if we were able to show that the small firms' complaints about high interest rates are the consequence of firms using expensive sources of credit when they are credit rationed, and not the consequence of banks offering only too high interest rates to small firms, we may come to a different conclusion about the need of establishing an interest rate cap.

Additionally, since one of the main objectives of investigating the determinants of financial constraints is to provide insights for public policy applications, we should be able to differentiate the effects that different instruments of public intervention could have on firms' access to bank credit and its cost. For instance, a certain policy could be more effective in reducing the cost of credit, and another policy could be more effective in improving the accessibility of formal credit. Hence, not differentiating between different types of constraints could reduce the applicability of the strategies and the accuracy of the evaluation of public efforts to reduce financial constraints.

This chapter addresses these issues by developing a new methodology to measure financing constraints, which identifies simultaneously credit rationed firms and firms facing restrictions in terms of unfavorable conditions, recognizing that these different situations can coexist. The solution proposed is to estimate different indicators that summarize the characteristics related to the types of constraints we want to measure using a multivariate technique called CATPCA, an adaptation of the original PCA method to analyze jointly both categorical and numerical variables.

This chapter performs this analysis on a database that provides direct firm information on all types of firms around the world, the World Business Environmental Survey (WBES) 2000, More specifically, I use this database to study which firm's characteristics and which credit market characteristics determine that firms have low access to bank credit or that they face unfavorable conditions of bank loans, such as high interest rates or short-term restrictions, while controlling for the different problems observed in the original variables of the survey.

By estimating two separated CATPCA analyses, I obtain the following results: With the first CATPCA model, I identify separately firms facing two different problems that cannot be differentiated with the original variables related to the access to credit. On the one hand I identify firms not having access to bank credit at all (credit rationed). On the other hand, I identify firms that do have access to bank credit, but only to short-term credit. Similarly, with the second CATPCA model I identify separately firms facing different problems that cannot be differentiated with the original variable related to high interest rates. On the one hand, I identify firms that have access to bank credit but that think that such credit is too expensive. On the other hand, I identify firms that do not have access to bank credit and are therefore forced to pay high interest rates to money lenders or suppliers. Finally, I can identify separately firms that do not demand any type of credit because they think that is too expensive.

With the results of the two CATPCA models I also obtain different variables (indicators) that measure the presence of these different constraints or problems that cannot be identified with the original variables. These new variables are the principal components (dimensions) that summarize the expected characteristics of each one of the cases described before.

After obtaining these different indicators, I run separate regressions with each one of them as dependent variables, while adjusting standard errors for a possible common country-level effect in the error term. Thanks to the statistical properties of the CATPCA methodology, such as the orthogonality of the components coming from each analysis, we can safely run regressions with each one of the new indicators as dependent variables without needing to control for whether the firm uses bank credit or not, or the source of credit they are using, since those problem are already taken into account in the construction of each variable.

The first result of the regression analysis indicates that access to bank credit clearly increases with firms' age and size and that the use of expensive sources of credit decreases as the firms grow and mature. However, as firms gain access to more bank credit, there is an initial increase in the perception of facing unfavourable conditions in the terms of the loans, such as high interest rates and maturity restrictions. This initial increase reach a maximum point, after which the restrictions related to those unfavourable conditions start decreasing, probably because of the decrease in informational asymmetries and because older and larger firms possess more assets that can be pledged as collateral, which at a certain point should have an effect on the actual interest rates paid.

Therefore the results indicate that there is an inverted u-shaped relationship between the perception of high interest rates of bank credit and the firm's age. This result has never been reported before for a measure of financial constraints or a measure of the perception of firms about too high interest rates. A similar evolution of the actual costs of credit has been reported by authors researching the effects of bank switching costs, as in Kim, Kristiansen and Vale (2008). Additionally, a similar inverted u-shaped relation is found between the indicator of scarcity of long-term financing and the age of the firm.

Another important result is that banking competition (less concentration) improves firm's access to credit but it does not seem to have a significant effect on the perception of a high cost of bank credit. Another novel result is that although banking concentration worsens the general access of firms to bank credit, there is some evidence that it may somehow increase the access in terms of short term credit. These results seem to give partial support to the structure-performance hypothesis of banking concentration, since I find a positive relation between the main indicator of access to credit and bank concentration, but no relation between the main indicator of the costs of bank credit and concentration.

It was also found that the perception of facing unfavourable conditions of bank credit is smaller in markets where regulation facilitates banks to seize collateral assets after bankruptcy. However, I did not find a significant effect on the indicators of credit rationing. This means that stronger creditor protection seems to improve financial restrictions by reducing the costs and improving the conditions of bank loans, but it does not seem to improve firms' access to bank credit.

The results also support the idea that firms in markets with more credit information sharing have more access to credit (less credit rationing), although it does not have an effect on the restrictions related to adverse conditions of bank loans. Additionally, more credit information sharing is found to increase the demand of credit from firms that think that credit is too expensive.

With respect to credit information sharing I also find different effects of the coverage of private bureaus versus public registries agencies. A higher coverage of private bureaus is found to improve firms' access to credit, but on the contrary, a higher coverage of public registry agencies is found to increase restrictions in terms of low access to long term credit, which can imply that a higher coverage of public registry agencies only helps improve the access to credit with short-term or adverse conditions. This difference could be important for policy makers trying to foster financial development by increasing the coverage of both private and public bureaus.

Finally, I find that credit rationing worsens with worst macroeconomic conditions. However, I do not find a significant effect on the indicators of high costs of bank credit or restrictions in the maturity. This means that access to credit seems to be pro-cyclical, but the perception of paying too high costs for credit is not affected by the economic cycle.

The rest of the chapter is organized as follows: Section 3.2 describes the data and summary statistics. Section 3.3 describes the methodology. Section 3.4 shows the main results and section 3.5 concludes.

3.2 Data and Summary Statistics

The main data source used is the World Business Environmental Survey (WBES) that is a survey run by the World Bank with the purpose of understanding better the constraints that hinder the development of private businesses.

Other variables also come from World Bank databases. Some variables come from the "Doing Business" database that provides objective measures of business regulations and their enforcement. The Doing Business indicators are comparable across 175 countries. They indicate the regulatory costs of business and can be used to analyze specific regulations that enhance or constrain investment, productivity, and growth.

The WBES firm-level dataset consists of firm survey responses of over 10,000 firms in 80 developed and developing countries. It provides information on firm size, age, employment, government ownership, foreign ownership, etc. The survey has a large number of questions on the business environment in which firms operate, including assessment of growth obstacles the firms face. The database also includes information on firm sales, industry, growth, financing patterns, and number of competitors.

The WBES includes different questions to firms that intend to state if they are facing financing constraints/obstacles for the operation and growth of the firm. More specifically they are asked how problematic different financing issues for their operation and growth are. Since I am interested in evaluating the determinants of credit rationing and costly external finance, the most important variables for the analysis are the following:

HINT: high interest rates
LTLOAN: access to long term loans

These variables are categorical (ordinal) variables that range from 1 to 4, where 1 corresponds to "no obstacle" and 4 to "major obstacle".

Next, there are other questions/variables in the WBES that include important information about the firms' financing patterns. Specifically, firms are asked to determine the share (percentage) of the firm's financing over the last year coming from each of the following sources:

- FN_RE: internal funds/Retained earnings
- BANK CREDIT: Domestic and Foreign commercial banks
- FN_FLY: Family/friends
- FN_MLDR: Money lenders
- FN_SHARE: equity, sale of stock
- FN_SCCR: supplier credit

In the analysis I merge the percentage of use of family/friends and moneylenders into one variable called informal:

- INFORMAL: moneylenders, family, friends, or informal sources of finance.

I also create a categorical variable BANK defining which firm uses bank credit and which firm does not:

- BANK, categorical variable that takes the value of 0 if firm does not use bank credit and 1 if firm uses credit from banks.

Variables defining firm characteristics included in the WBES are:

- SIZE_F, a categorical variable measuring the size of the firm according to the number of employees. It takes the value of 1 for small firms, 2 for medium size firms and 3 for large firms. We define firm size according to the number of employees. A firm is small if it has between 5 to 50 employees, medium-size between 51 to 500 employees and large with more than 500 employees.
- AGE, the number of years since the firm was originally established.
- FOREIGN, a dummy variable that takes the value of 1 if the firm is from a foreign country.
- GOVERNMENT, a dummy variable that takes the value of 1 if the firm is publicly owned.
- SERVICE, a dummy variable that takes the value of 1 if the firm's main business is in the service sector.
- MANUFACTR, a dummy variable that takes the value of 1 if the firm's main business is in the manufacturing sector.
- SMALL, a dummy variable that takes the value of 1 if the firm has between 5 and 50 employees,
- MEDIUM, a dummy variable that takes the value of 1 if the firm has between 50 and 500 employees.
- LARGE, a dummy variable that takes the value of 1 if the firm has more than 500 employees.

The variable AGECAT was categorized according to the following rule:

- AGECAT, a categorical variable that takes the value of 1 if the firm is younger than 5 years, 2 if the firm is between 5 and 20 years old and 3 if the firm is older than 20 years.

The country-level variables that we include in the regression analysis are:

- CONCENTR, a continuous variable measuring the concentration of the banking sector, measured by the assets of the three largest banks as share of all commercial banks in the system.
- LEGAL_RIGHT, it is an index of the strength of legal rights. It measures the degree to which collateral and bankruptcy laws protect the rights of borrowers and lenders and thus facilitate lending. The index includes 7 aspects related to legal rights in collateral law and 3 aspects in bankruptcy law.

- CRED_INFORM, it is an index of the depth of credit information and it measures rules affecting the scope, accessibility and quality of credit information available through either public or private credit registries.
- PUB_REGSTR, It is and indicator of the public credit registry coverage. It reports the number of individuals and firms listed in a public credit registry with current information on repayment history, unpaid debts or credit outstanding.
- PRIV_BUREAU, it is an indicator of the private credit bureau coverage and reports the number of individuals or firms listed by a private credit bureau with current information on repayment history, unpaid debts or credit outstanding.
- GROWTH, the average rate of growth of GDP in the previous 3 years to the survey.

A complete description of these variables and other relevant variables included in the survey is provided in Appendix A.I.

In the following tables we can see the average of the main variables defining firm characteristics according to firms' size and age, where SIZE_F and AGECAT are categorical variables.

Here it is important to highlight that smaller and younger firms (SMALL and AGECAT=1) are the ones with the highest mean in the variable LTLOAN but the highest mean for the variable HINT corresponds to MEDIUM firms and AGECAT=2 and the lowest correspond to the largest and oldest firms. Intuitively, smaller and younger firms face higher asymmetric information problems and higher relative costs of screening and monitoring, and thus, it may be expected that young and small firms are the ones who complain the most about high interest rates. But instead, medium size and age firms are the ones that actually complain the most.

It is also important to highlight that younger and smaller firms are the ones with the lowest use of bank credit as it can be seen in Table 1 and Table 2. At the same time they are the ones with the highest use of informal sources (family and money lenders). This is important because it may indicate that younger and smaller firms are not the ones who complain the most about the cost of bank credit, but that their complaints about high interest may come from their use of expensive sources of credit when they lack access to formal credit from banks.

Table 1 Financing Obstacles according to firm characteristics

	AGE	SIZE	HINT	LTLOAN
AGECAT=1	3.12	1.50	3.27	2.85
AGECAT=2	8.71	1.65	3.28	2.73
AGECAT=3	41.10	2.09	3.24	2.50
SMALL	9.93		3.26	2.77
MEDIUM	19.69		3.31	2.69
LARGE	31.94		3.20	2.54
MANUFACT	20.62	1.95	3.30	2.72
SERVICE	15.36	1.58	3.16	2.57
GOVMNT	25.86	2.21	3.18	2.58
FOREIGN	22.92	2.16	3.03	2.41
BANK=0	14.78	1.61	3.21	2.72
BANK=1	21.52	1.92	3.34	2.65
TOTAL	17.61	1.74	3.27	2.69

Table 2 Financing Sources according to firm characteristics

	Bank Credit	Retained Earnings	Sale of Stock	Supplier Credit	Informal
AGECAT=1	10.9	61.7	4.0	5.7	12.8
AGECAT=2	17.7	56.0	5.8	6.8	7.9
AGECAT=3	24.2	51.1	4.9	8.2	4.2
SMALL	11.8	58.2	5.1	6.0	14.8
MEDIUM	19.0	56.7	4.6	8.0	4.4
LARGE	28.4	50.0	4.8	6.3	2.0
MANUFACT	21.4	55.3	4.6	7.1	6.6
SERVICE	15.2	56.5	5.4	6.1	10.1
GOVMNT	15.7	57.0	4.9	5.2	2.0
FOREIGN	24.3	51.7	7.2	7.2	3.4
BANK=0	0.0	71.7	5.0	5.6	10.9
BANK=1	41.8	34.9	4.7	8.7	4.8
TOTAL	17.5	56.3	4.9	6.9	8.3

Table 3 Financing Sources according to financing obstacles

		Bank Credit	Retained Earnings	Sale of Stock	Supplier Credit	Informal
HINT	4	17.6	55.4	4.1	7.8	9.1
	3	21.6	52.1	6.7	6.9	6.7
	2	18.6	56.9	6.2	5.5	6.5
	1	13.1	64.0	4.3	4.3	7.6
LTLOAN	4	15.7	55.8	3.1	8.6	10.7
	3	20.0	51.7	5.8	7.4	8.7
	2	22.7	50.0	8.3	6.6	6.3
	1	19.1	59.5	5.3	5.2	5.4
Total		18.4	55.4	4.9	7.1	8.2

Table 4 Correlations between Financing Obstacles and financing sources

	HINT	LTLOAN	Use of bank credit (=1)	Bank Credit	Supplier Credit	Informal	Sale of Stock	Retained Earnings
HINT	1							
LTLOAN	0.4655	1						
Use of bank credit (=1)	0.0854	-0.0031	1					
Bank Credit	0.0641	-0.0234	0.7303	1				
Supplier Credit	0.0817	0.0892	0.1041	0.0018	1			
Informal	0.0455	0.1095	-0.1152	-0.1165	-0.0479	1		
Sale of Stock	-0.0096	-0.0534	0.0035	-0.0458	-0.0133	-0.0519	1	
Retained Earnings	0.0054	0.0069	-0.3603	-0.4413	-0.2744	-0.2997	-0.2323	1

Table 5 Correlations between firm characteristics, financing obstacles and financing sources

	SIZE	AGE
SIZE	1	
AGE	0.3071	1
HINT	-0.0236	-0.0555
LTLOAN	-0.059	-0.1069
Use of bank credit (=1)	0.206	0.1193
Bank Credit	0.1931	0.1125
Supplier Credit	0.0358	0.0178
Informal	-0.225	-0.1109
Sale of Stock	0.0055	0.0208
Retained Earnings	-0.0361	-0.0522

It is also important to emphasize some characteristics of the variables HINT and LTLOAN that do not fit into the a priori expectations of these types of constraints. First of all, if a variable is expected to be a good measure of credit rationing (low access to credit) we would expect it to have a large negative correlation with the use of bank credit. However, LTLOAN shows almost zero correlation (-0.0031) with the dummy variable BANK and a very low correlation with the share of bank credit in the financing sources (-0.0234) (Table 4).

Although we may expect that credit constrained firms are forced to use internal funds when they lack access to credit, LTLOAN shows also a near zero correlation with the use of retained earnings (0.0054) (Table 4). The sources of credit with the highest correlation with LTLOAN are “Informal” (0.1092) and Supplier Credit (0.0892).

The variable HINT shows a positive and similar correlation with the use of bank credit (0.0854) and supplier credit (0.0817) and a positive correlation with the use of informal sources (0.0455) and almost no correlation with the use of retained earnings (0.0054).

The positive correlation with the different financing source indicates that if we want to determine the determinants of the variable HINT, it is important to control for the type of credit source that the firm is using, since the determinants could be different if a firm complains about high interest rates of bank credit or high interest rates of informal sources or supplier credit. A similar situation applies for the variable LTLOAN, since the determinants of this variable could be different if a firm does not have access to credit at all, or if a firm have access to credit, but only to short-term credit.

For instance, we can see that although both HINT and LTLOAN show negative correlations with the size and age of the firm, the correlations are not high, although are higher for LTLOAN. These low correlations could be the consequence that these two variables are related to the use (or not) of different sources of credit.

Finally, the low correlation of HINT with the use of retained earnings is also surprising, since we may also expect that if a firm has to pay a high interest for bank crediit, it may want to use internal funds instead.

3.3 Methodology

3.3.1 CATPCA Analysis (Financial Constraints Indexes)

I chose to use a dimension-reducing multivariate technique analysis for several reasons. First, the WBES do not include a direct measure of credit rationing or the bank loans' interest rates paid by each firm. However, in the survey there are many variables that could partially indicate the presence of these constraints.

In principle, the variables in the survey that are aimed to indicate the presence of constraints in terms of low access to credit and high costs of credit are LTLOAN and HINT. However, these variables present some important shortcomings.

On the one hand, the variable LTLOAN only asks whether the access to long-term loans is an obstacle for the firm. But many firms may not have access to bank credit at all and others may only have access to short term loans (maturity restrictions).

Additionally, for determining if a firm is facing credit rationing we may use the dummy variable BANK that indicates if the firm is using bank credit or not. However, as we do not observe the firm's demand for funds we cannot conclude if the absence of bank credit is just because the firm does not demand any credit.

On the other hand, the variable HINT has also different important problems. HINT only says whether high interest rates are an obstacle for the firm, but we do not observe the actual interest rate paid by each firm for their loans; neither can we observe the type of credit that the firm is actually complaining about. Therefore, if a firm complains that high interest rates are an obstacle we cannot know if it is because a bank is charging the firm a high interest rate, or if a firm that does not have access to bank credit is forced to pay high interest rates to money lenders or suppliers.

Furthermore, the question about the perception of high interest rates do not differentiate between firms that are actually paying high interest rates for the credit they are using and firms that have do not demand any type of credit because it is too expensive, and they use internal funds to partially cover their investment needs.

In summary, the way the questions were phrased in the survey limits the interpretation that we could give them and the conclusions we can extract if we want to estimate regression models to determine their main determinants. It is clear that if we run regressions with HINT and LTLOAN as dependent variables we would need to control for the type of credit they are using (or the credit they are not using). However, simply introducing the different sources of credit as independent variables is not an appropriate solution. Clearly, we would have an endogeneity problem, since the source of funds that the firm uses is simultaneously determined to the perception of its cost and to the problem of having or not having access to credit.

One option would be to estimate a system of equations with all these variables as endogenous ones. For instance we could run a two stage regression with instrumental variables to control for the endogeneity of HINT, bank credit and other sources of funds. However we would need to find instruments that are correlated with the sources of credit (bank credit, informal sources or supplier credit) but that, at the same time, are uncorrelated with the perception of paying high interest rates (HINT).

Alternatively, we can estimate different indicators that summarize the characteristics related to the type of constraints we want to measure. In other words, we think that there exist some latent or hidden variables (credit rationing or high costs of different types of credit) that are affecting simultaneously all the firms' financing variables but that are not correctly captured by the original variables. Principal component analysis (PCA) allows us to construct a measure of these hidden or latent variables that we are not directly observing.

The specific multivariate technique that I am using is called Categorical Principal Component Analysis (CATPCA), an adaptation of the original method to analyze jointly both categorical and numerical variables. This procedure simultaneously quantifies the categorical variables at the same time as reducing the dimensionality of the data. The objective of the CATPCA is the reduction of an original set of variables into a smaller set of non-correlated components that represent most of the information found in the original variables. The categorical variables are quantified optimally on scales of the specified dimensionality, using optimal scaling as in multiple correspondence analysis (see, for example, Greenacre (1993)).

The multivariate technique used here allows us to do this easily, because it allows us to combine categorical (e.g. LTLOAN, HINT, BANK) and continuous variables (e.g. FN_SCCR and INFORMAL).

Using CATPCA allows us to accomplish the following objectives:

1. We can identify separately firms facing two different problems that cannot be differentiated with the original LTLOAN variable. On the one hand we can identify firms not having access to bank credit at all (credit rationed). On the other hand, we can identify firms that do have access to bank credit, but only to short-term credit. We can obtain an indicator for each different problem.
2. We can identify separately firms facing two different problems that cannot be differentiated with the original HINT variable. On the one hand, we can identify firms that have access to bank credit but that think that such credit is too expensive (and thus complain about high interest rates). On the other hand, we can identify firms that do not have access to bank credit and are thus forced to pay high interest rates to money lenders or suppliers (and thus complain about high interest rates). Additionally, we can construct a different indicator of each one of the previous problems.
3. We can also separate firms that are actually paying high interest rates for credit (the problems in point 2) from those that do not demand any type of credit because they think that is too expensive. We can also get an indicator for this type of situation.

In order to achieve the three objectives described I run two separated CATPCA models. The first model uses the variables LTLOAN and BANK to achieve the first objective. The results from the first model and the new indicators should have the following structure and characteristics:

$$COMPONENT\,(1) = a_1 LTLOAN + a_2 BANK$$
$$COMPONENT\,(2) = b_1 LTLOAN + b_2 BANK$$

- $a_1 > 0$ and $a_2 < 0$, The first component should measure whether a firm has no access to credit at all. Thus, the index will be called RATIONED and it should be higher the higher the complaints about access to long-term loans ($a_1 > 0$), showing that firms demand more funds than the ones actually used. Additionally, the index should be higher when the firm does not use bank credit, i.e. when the dummy BANK=0 (or lower when BANK=1) ($a_2 < 0$).

- $b_1 > 0$ and $b_2 > 0$, The second component should measure whether a firm is rationed in terms of not having access to long-term finance, meanwhile having access to short-term bank credit. Thus, the component will be called RATIONED_LT and it should be higher the higher the complaints about access to long-term loans ($b_1 > 0$). Additionally, the index should be higher when the firm uses bank credit, i.e. when the dummy BANK=1 ($b_2 > 0$), showing that the firm has access to bank loans, but only to short-term ones.

If RATIONED and RATIONED_LT show the described descriptions they could be used as two new indicators of the two types of restrictions described in the first objective. Since they are by construction totally uncorrelated to each other (orthogonal) we can safely run a regression with each one as dependent variables without needing to control for whether the firm uses or not bank credit, since this problem is already taken into account in the construction of each variable.

The second CATPCA model uses the variables HINT, BANK CREDIT, INFORMAL and FN_SCCR (supplier credit) in order to achieve the second and third objectives. The results from the model and the new indicators should have the following structure and characteristics:

$$COMPONENT\,(1) = a_1 HINT + a_2 BANK + a_3 (FN\_SCCR + INFORMAL)$$
$$COMPONENT\,(2) = b_1 HINT + b_2 BANK + b_3 (FN\_SCCR + INFORMAL)$$
$$COMPONENT\,(3) = c_1 HINT + c_2 BANK + c_3 (FN\_SCCR + INFORMAL)$$

- $a_1 > 0, a_2 > 0$ and $a_3 \leq 0$, The first component is supposed to be a measure of firms complaining only about the high interest rates of bank credit with no relation to the use of informal sources or supplier credit, and therefore, it will be called HINT_BANK. It should be higher the higher the complaints about high interest rates, showing that firms face bad conditions in the use of credit ($a_1 > 0$). This indicator should also be higher if the firm is using bank credit ($a_2 > 0$) assuring that the complaints are related to the use of credit from banks. Finally, the indicator should be lower or uncorrelated to the use of supplier credit of informal sources ($a_3 \leq 0$), guaranteeing that the firms that score high in this index do so because they face high costs of *bank* credit and not because they use other costly types of credit such as money lenders.

- $b_1 > 0, b_2 < 0$ and $b_3 > 0$, A second component of the model could also be a measure of low access to bank credit. It is supposed to indicate a situation in which a firm does not have access to bank loans and therefore complains about the high interest rates of other more expensive sources of credit. Therefore, it will be called HINT_NOBANK. Such indicator should be higher the higher the complaints about high interest rates ($b_1 > 0$), but such complaints should be associated to the use of sources credit different from bank loans. Therefore, it should be lower if the firm is using bank credit ($b_2 < 0$) (low access to credit) and it should be higher the higher the use of supplier credit or informal sources ($b_3 > 0$).
- $c_1 > 0, c_2 < 0$ and $c_3 < 0$, The last component of the model is supposed to capture the problem of those firms who choose to not use bank credit or any other type of credit because they think that they are too expensive. Therefore, it will be called HINT_NOCREDIT. This indicator should be higher the higher the complaints about high interest rates ($c_1 > 0$) and it should be lower if the firm is using bank credit ($c_2 < 0$), supplier credit or informal sources ($c_3 < 0$).

Similar to the previous case, since these new indicators (HINT_BANK, HINT_NOBANK and HINT_NOCREDIT) are by construction totally uncorrelated to each other, we can safely run regressions with each one as dependent variables without needing to control for the source of credit they are using, since such problem is already taken into account in the construction of each variable.

a_i, b_i and c_i re the weights of variable *i* in the respective indicator. These weights are given by the score of each variable in the adequate component of the CATPCA model.

PCA transforms a set of correlated variables (*x's*) into a set of uncorrelated components (*y's*). The principal components are linear combinations of the *x's*. Each component is a weighted sum of the *x's*, and in the case of CATPCA, each component is a weighted sum of each category in every variable.

In PCA, if we call the weight given to variable *i* on component *j* as a_{ij}, then the relative sizes of the a_{ij} reflect the relative contribution made by each variable to the component. To interpret the component, we examine the pattern in the a_{ij} values for that component. In the case of CATPCA we have weights not only for each variable, but for each category of every variable. Thus, it is easier to interpret the component loadings. The component loadings can be interpreted as the correlation coefficient between variable *i* and component *j*.

The total number of components is equal to the number of variables (*x's*) and, as the *x's* are standardized, the sum of the variances of all the variables is also equal to the number of variables.

When performing CATPCA I need to categorize the continuous (numerical) variables such as FN_SCCR and Informal. When doing this, I took into account the statistical distribution of the different variables and I used the same categorization method for different variables. All the continuous variables were categorized using a grouping by equal intervals, or in other words the variables were codified in categories defined by intervals of the same length.

3.3.2 Regression Analysis: Determinants of Different Financial Constraints

After obtaining the different indicators, I run the next basic regressions, adjusting standard errors for a possible common country-level effect in the error term[3]:

INDICATOR(i) = *f*(SMALL, LARGE, AGE, SQRT_AGE, LEGAL_RIGHTS, CREDIT_INFORMATION, CONCENTRATION, GROWTH, MANUFACTURING, SERVICE, FOREIGN, GOVERMENT)

INDICATOR(i) corresponds to each one of the previously defined variables, i.e. RATIONED, RATIONED_LT, HINT_BANK, HINT_NOBANK and HINT_NOCREDIT.
Since RATIONED and HINT_NOBANK are both related to firms facing low access to bank credit, we may expect that they have similar characteristics and determinants. In contrast, HINT_BANK and RATIONED_LT are both related to firms facing restrictions in the conditions of their bank loans, such as the cost of the loan or its maturity. Thus, we may also expect that these two variables show similar characteristics and determinants. Firms scoring high in the HINT_NOCREDIT indicator may share some characteristics and determinants with the other indicators.

According to theory, we may expect that credit rationing decreases with the firm's size and age, because older and larger firms are less informational opaque and because older firms may have longer and wider relationship to creditors. Therefore, in principle, we may expect a negative relationship between the RATIONED and HINT_NOBANK indicators with the size and age of the firms, because as firms grow and have more access to bank credit (lower RATIONED) they also could rely less on expensive sources of credit, reducing HINT_NOBANK.

On the other hand, the costs of bank credit may be lower for larger and older firms because they possess more assets that can be pledge as collateral and because of less informational asymmetries according for instance to Winker (1999), Petersen and Rajan (1994), and Cooley and Quadrini (2001).

However, if younger and smaller firms are rationed from bank credit, they cannot face strong restrictions in terms of its costs or maturity. As younger firms grow older they become less informational opaque for banks, and therefore, they can access more bank credit and complain more about the cost and maturity of it. Additionally, as firms become older and start

[3] Robust estimator of variance allowing for country clustering.

receiving credit and informational asymmetries become less important, banks should be able to price credit better and therefore, the cost of bank loans should start decreasing with age at some point in time.

In summary, we may find a non linear effect of AGE on the HINT_BANK and RATIONED_LT indicators, which are the ones related to the conditions of bank loans, and related to firms which already have access to bank credit. Therefore, I also test for the possibility that the effect of age on all the indicators is non linear. For doing this I include the square root of age in the regressors, SQROOT_AGE.

We are highly interested in the effect of banking concentration on the credit market (CONCENTR). It is usually argued that banking concentration should decrease the availability of credit and should increase the price of credit, but as explained in the introduction to this chapter, this effect is controversial in the literature.

According to the conflicting theories of banking concentration, we expect that if the structure-performance hypothesis (market power inefficiencies) is correct, we should find a positive relation between banking concentration and credit rationing and the costs of bank credit, and therefore, a positive relation between concentration and all the indicators. Meanwhile, if the information-based hypothesis is correct, we should find a negative relation between the concentration and the indicators of access to credit (RATIONED and HINT_NOBANK) and a positive or nonlinear relation between the indicators of the costs of bank credit (HINT_BANK and RATIONED_LT) and banking concentration. However, to the best of my knowledge, all the studies that have investigated the relation between access to credit and interest rates with the concentration in the banking industry have used the observed interest rates and not the perception of the firms about whether interest rates are too high or to low.

Restrictions in terms of credit rationing and high costs of bank debt should decrease with a better legal environment, i.e. the easier the bank can seize firm's assets the more valuable is collateral to bank, and therefore, the less costly should be formal credit. Therefore, I expect a negative relationship between all indexes and the variable LEGAL_RIGHTS, see for instance Beck et al (2004), Manove et al (2001) and Berger and Udell (1998).

Credit rationing should decrease with more credit information sharing, i.e. if banks have more and better information about the firm's quality and risk (see for instance, Pagano and Japelli (1993)). Thus, RATIONED and HINT_NOBANK should be negative related to CREDIT_INFORM.

On the contrary the effect of credit information sharing on the costs of bank credit may be ambiguous and similar to the effect of the age of the firm. On the one hand, with more credit information banks should be able to better price the loan's interest rate and informational asymmetries should be lower on a market with better information sharing among creditors. On the other hand, more credit information allows the firm to have more access to formal credit and therefore, and in general, to more costly credit. Therefore, the effect of credit information on the cost of credit depends on which effect dominates.

If the effect of more credit information is to increase the access to more costly credit, then HIGHCOST index should be positively related to CRED_INFORM. If the effect of more credit information sharing is to reduce a possible overpricing of the loan's interest rate then the effect of CRED_INFORM on HIGHCOST should be negative.

Finally, in the basic regression specification, I control for the effect of the macroeconomic environment in the respective country (GROWTH). If financial constraints are counter-cyclical then we expect a negative effect of GROWTH on both indexes.

I also run Ordinal Probit Regressions of the same equations for the CREDRAT and HIGH-COSTS indexes, but using the original survey variables HINT and LTLOAN as dependent variables, in order to compare the results and the added value of the new variables with respect to the original one.

LTLOAN = *f*(SMALL, MEDIUM, AGE, SQRT_AGE, LEGAL_RIGHT, CRED_INFORM, CONCENTR, GROWTH, MANUFACTURING, SERVICE, FOREIGN, GOVERMENT)

HINT = *f*(SMALL, MEDIUM, AGE, SQRT_AGE, LEGAL_RIGHT, CRED_INFORM, CONCENTR, GROWTH, MANUFACTURING, SERVICE, FOREIGN, GOVERMENT)

Additionally, as a way to further show the added value of the results with respect to previous studies, I run the same ordinal probit regressions as in Beck et al (2006) using the variables LTLOAN and HINT as dependent variables, and using country dummy variables to control for common country characteristics instead of the specific country variables used in the previous analysis. This approach does not allow controlling for specific country variables such as concentration, credit information sharing or creditors' protection, given that these variables are constant across firms in the same country, and therefore, country variables would be almost perfectly collinear with the country dummy variables.

3.4 Estimation Results

3.4.1 CATPCA Analysis

The estimation results for the CATPCA are shown in Table 6 to Table 9. In Table 6 and Table 7 I show the results of the CATPCA model for the variables LTLOAN and BANK(1,0). In Table 8 and Table 9 I show the results of the CATPCA model for the variables HINT, BANK(1,0) and (FN_SCCR+INFORMAL).

According to the correlations of the components with the original variables, firms with the highest score in the first component of the first CATPCA model (Table 7) are the firms that complain more about having problems to access long-term loans (high value of LTLOAN) and firms that at the same time do not use any bank credit (high negative correlation with BANK). These characteristics show that they are demanding more bank credit than the actual received. Firms scoring the lowest in this component are firms that do use bank credit and have no complaints about access to long term loans. Therefore, I use this first component as a measure of credit rationing, and I call it as RATIONED.

Firms with the highest scores in the second component of the first model (Table 7) are firms that complain about having problems to access long-term loans (high value of LTLOAN) but also that, differently to the previous case, do use bank credit (high positive correlation with BANK). These characteristics show that although they are not totally rationed from credit, they demand loans with different conditions, more specifically, with longer maturities. In the other extreme of the component we can find firms that do not use bank credit but that do not complain about needing more. I call this component RATIONED_LT.

Firms with the highest score in the first component of the second CATPCA model (Table 9) are firms that have the highest use of supplier credit and informal sources of credit (family and money lenders) as this is the component with the highest correlation with these variables. On the contrary they have a very low use of bank credit (high negative correlation with BANK), and at the same time they have strong complaints about high interest rates.

The high positive correlation to Informal and supplier credit together with the high negative correlation to BANK indicate that the complaints about high interest rates of the firms with a high score in this component come from the use of expensive sources of credit different from bank credit. We can think of this component as another measure of access to bank credit, since firms scoring high in this component are firms that do not have access to bank credit and are therefore forced to use alternative but more expensive sources of funds. I call this component HINT_NOBANK.

Firms with the highest score in the second component of the second CATPCA model (Table 9) are firms that have the strongest complaints about high interest rates (highest correlation to HINT). At the same time, this component is positively correlated to the use of bank credit (BANK) and show an almost zero correlation to the use of informal sources or supplier credit (0.03). These three characteristics indicate that this component is measuring the complaints about high interest rates of firms that are actually using bank credit to finance their investment. Firms with a low score in this component are firms with a low use of bank credit, but that do not complain about the interest rates. Therefore, I called this component HINT_BANK.

Finally, firms with the highest score in the third component of the second CATPCA model (Table 9) are firms that have a very low use of all the possible sources of credit, which can be seen from the high negative correlation with BANK and Informal sources and supplier credit. Therefore, this component is indicating those firms that have a very low demand of any type of credit because of their high interest rates, which is an obstacle for them. Therefore, I call this component as HINT_NOCREDIT.

The new variable RATIONED accounts for 51.5% of the total variance of the original variables LTLOAN and BANK included in the first CATPCA model. The new variable RATIONED_LT accounts for 48.5% of the total variance of the two original variables (Table 6). Therefore, both variables account for a similar amount of information. Since I am using the two components of the analysis I am not losing any information from the original variables. The correlations of the original variables with the two components of the first CATPCA model are shown in Figure 1.

The new variable HINT_NOBANK accounts for 37.4% of the total variation of the original variables included in the second CATPCA model. The new variable HINT_BANK accounts for 34.4% and the new variable HINT_NOCREDIT for 28.7% (Table 8). The similar proportions of variance explained by each component indicate that the three new variables are important in understanding the total variation and information contained in the original variables. The correlations of the original variables with the two components of the second CATPCA model are shown in Figure 2.

Table 6 Summary of first CATPCA Results: Eigen Values and Percentage of Variance of Each Component

Component	Eigen Values	% of Variance	% Accumulated
1	1.030	51.48	51.48
2	0.970	48.52	100.00

Table 7 Summary of first CATPCA results: loadings of the principal components

	1 RATIONED	2 RATIONED_LT
LTLOAN	0.7175	0.6965
BANK (1,0)	-0.7175	0.6965

Table 8 Summary of second CATPCA Results: Eigen Values and Percentage of Variance of Each Component

Component	Eigen Values	% of Variance	% Accumulated
1	1.1065	36.88	37.37
2	1.033	34.43	71.8
3	0.8605	28.68	100

Table 9 Summary of second CATPCA results: loadings of the principal components

	1 HINT_NOBANK	2 HINT_BANK	3 HINT_NOCREDIT
HINT	0.4483	0.7536	0.4808
BANK CREDIT	-0.5279	0.6815	-0.5069
INFORMAL + SUPPLIER	0.7918	0.0277	-0.6102

Figure 1 First and Second Components of first CATPCA model

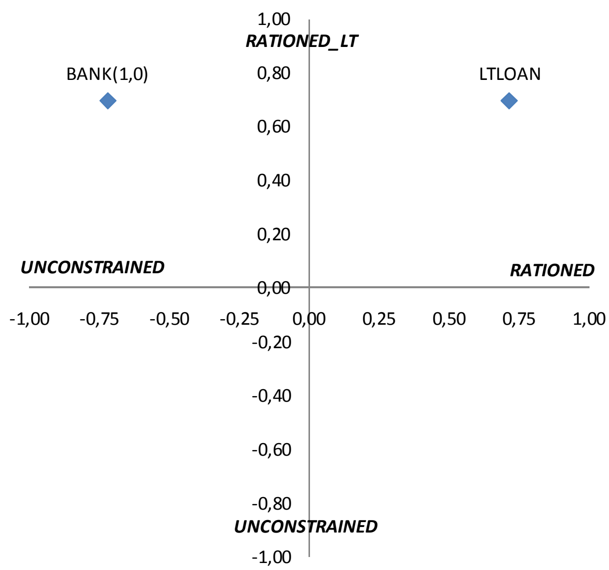
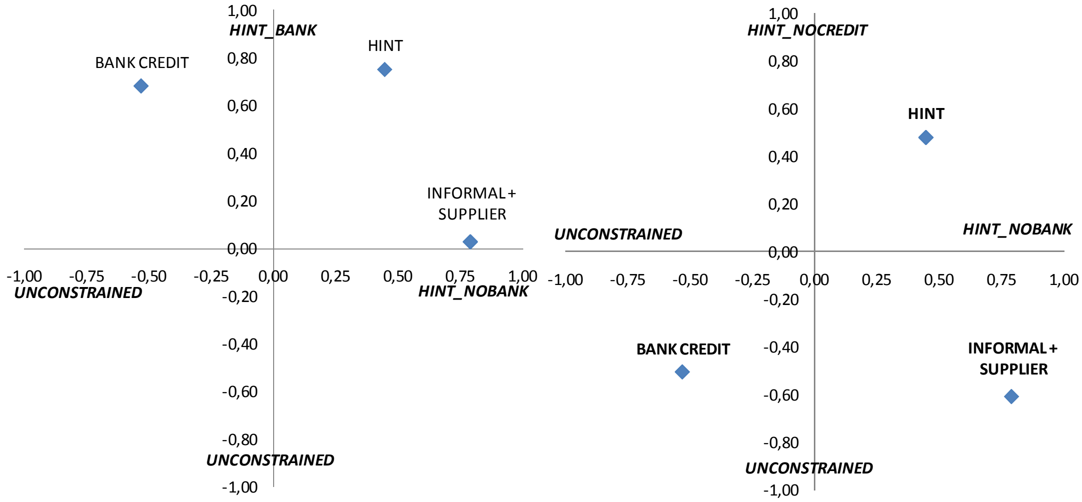


Figure 2 First vs. Second and First vs. Third Components of second CATPCA model



The new variables show some characteristics that help clarifying the inconsistent characteristics of the original variables HINT and LTLOAN as measures of credit rationing and high interest rates of bank credit. First of all, as seen in Table 10 and Table 11, RATIONED and HINT_NOBANK are highly correlated to each other (0.463), as they are both different measures of low access to bank credit. Both variables have high and negative correlations to the use of bank credit (BANK and BANK CREDIT) and both are positively correlated to LTLOAN as well.

Firms scoring high in these two indicators are firms that clearly do not have access to formal bank credit. The difference among the two is that firms scoring high in RATIONED are able to use retained earnings to replace bank credit and the ones scoring high in HINT_NOBANK deal with the lack of bank credit by replacing it with informal sources and supplier credit. These can be seen in the high positive correlation between RATIONED and the use of retained earnings (0.2982) and the correlations between HINT_NOBANK and retained earnings (-0.127), informal sources (0.583) and supplier credit (0.478).

In the same fashion, we can see in Table 12 that the firms in the quartile with the highest score in RATIONED use mainly retained earnings (69.2%), informal sources (13.6%) and supplier credit (7.12%) to finance their investments. These firms basically do not use bank credit at all (0%). On the other hand, the firms in the quartile with the highest score of HINT_NOBANK have one of the lowest uses of retained earnings (31.3%), meanwhile they show the highest use of informal sources (29.62%) and supplier credit (21.85%). According to their lack of internal funds, they use some bank credit, but at a low level (8.63%).

Similarly, the indicators HINT_BANK and RATIONED_LT are also positively correlated (0.614) since they are both measures of firms facing restrictions in terms of bad conditions in their bank loans (high cost and short maturity). Both indicators are, however, positively correlated to the use of bank credit.

We can see in Table 12 that the firms in the quartile with the highest score of HINT_BANK are the ones with the lowest use of retained earnings (23.3%) and the highest use of credit from banks (54.08%). This characteristic could be a sign that they are the ones more highly leveraged, which could be the reason of the high interest rates that they complain about (76% of their investment is financed with different sources of credit). Likewise, firms in the quartile with the highest score in RATIONED_LT have a very low use of retained earnings (35.8%) and a high use of bank credit (39.9%).

Finally, it is clear that firms scoring high in the HINT_NOCREDIT indicator are firms that basically rely on internal funds because they find interest rates on the different types of credit too expensive. Firms in the quartile with the highest score in this indicator have the highest use of retained earnings (87.4%) and the sale of stock shares (5%), meanwhile they use no bank credit (0.1%), supplier credit (0.1%) or informal sources (0%).

Table 10 Correlation between firm characteristics, original financing obstacles and the new indicators

	SIZE	AGE	HINT	LTLOAN	rationed	rationed_lt	hint_nobank	hint_bank	hint_nocredit
SIZE	1								
AGE	0.3059	1							
HINT	-0.0136	-0.0515	1						
LTLOAN	-0.0614	-0.1132	0.4394	1					
rationed	-0.1877	-0.1699	0.2623	0.7148	1				
rationed_lt	0.1042	0.012	0.3599	0.6987	-0.0009	1			
hint_nobank	-0.2228	-0.1469	0.4467	0.3049	0.4627	-0.0374	1		
hint_bank	0.1238	0.0462	0.7519	0.2869	-0.1992	0.614	-0.0053	1	
hint_nocredit	-0.0157	-0.0429	0.4733	0.1751	0.4271	-0.1866	0.0023	-0.0117	1

Table 11 Correlation between new indicators and financing sources

	rationed	rationed_lt	hint_nobank	hint_bank	hint_nocredit
Use of bank credit (=1)	-0.7153	0.6994	-0.3567	0.5716	-0.4356
Bank Credit	-0.5553	0.4897	-0.5328	0.6699	-0.484
Supplier Credit	-0.0017	0.122	0.4779	0.066	-0.3993
Informal	0.1689	-0.0251	0.5827	-0.0467	-0.3828
Sale of Stock	-0.032	-0.0565	-0.0224	-0.065	0.0705
Retained Earnings	0.2982	-0.3525	-0.1268	-0.4235	0.6553

Table 12 Financing sources according to quartiles of new indicators

		Bank Credit	Retained Earnings	Sale of Stock	Supplier Credit	Informal
Rationed	4	0.0	69.2	3.3	7.1	13.6
	3	20.3	51.4	5.1	8.4	8.7
	2	0.0	77.3	5.7	4.0	7.0
	1	43.4	34.5	5.6	7.6	3.8
Rationed_lt	4	39.9	35.8	3.6	9.9	5.8
	3	9.8	60.7	4.3	7.3	11.3
	2	26.3	48.2	5.5	6.7	7.1
	1	0.0	75.3	6.5	4.2	7.6
hint_nobank	4	8.6	31.3	4.0	21.9	29.6
	3	1.7	84.4	4.9	1.1	0.6
	2	14.2	67.0	6.5	3.2	1.6
	1	47.9	41.3	4.0	1.3	0.4
hint_bank	4	54.1	23.3	3.2	10.2	5.4
	3	10.1	44.1	4.1	13.4	22.3
	2	6.4	75.1	5.4	3.2	3.4
	1	2.6	73.8	6.4	3.5	6.0
hint_nocredit	4	0.1	87.4	4.9	0.1	0.0
	3	7.8	72.4	6.6	2.8	2.4
	2	31.3	48.1	5.1	5.4	4.4
	1	31.6	17.5	2.7	19.4	25.7
Total		18.0	55.9	4.9	7.0	8.2

Table 13 Firm characteristics and original variables according to quartiles of new indicators:

		SIZE	AGE	HINT	LTLOAN
Rationed	4	1.57	11.90	3.67	4.00
	3	1.77	17.44	3.54	3.32
	2	1.71	18.89	2.46	1.00
	1	1.94	23.08	3.10	1.86
Rationed_lt	4	1.93	19.73	3.65	3.66
	3	1.63	14.28	3.57	3.53
	2	1.82	20.56	3.04	1.85
	1	1.69	18.52	2.65	1.27
hint_nobank	4	1.55	14.29	3.59	3.09
	3	1.70	14.75	3.95	3.04
	2	1.81	18.93	3.05	2.49
	1	1.93	22.89	2.46	2.13
hint_bank	4	1.93	21.47	3.76	2.88
	3	1.64	16.29	3.79	3.07
	2	1.72	15.32	3.63	2.86
	1	1.69	17.93	1.91	1.96
hint_nocredit	4	1.70	14.70	4.00	3.06
	3	1.77	17.99	3.33	2.70
	2	1.80	20.17	2.94	2.45
	1	1.71	17.59	2.85	2.59
Total	Total	1.75	17.68	3.27	2.69

Table 14 New indicators according to firms' characteristics:

	rationed	Rationed_lt	hint_nobank	hint_bank	hint_nocredit
AGECAT=1	0.30	-0.12	0.20	-0.16	0.09
AGECAT=2	0.00	0.05	-0.01	0.02	0.00
AGECAT=3	-0.28	0.08	-0.20	0.14	-0.09
SMALL	0.20	-0.12	0.23	-0.14	-0.02
MEDIUM	-0.06	0.05	-0.08	0.06	0.06
LARGE	-0.33	0.16	-0.40	0.21	-0.11
BANK=0	0.63	-0.62	0.31	-0.49	0.38
BANK=1	-0.81	0.79	-0.42	0.66	-0.50
MANUFACT	-0.06	0.10	-0.09	0.11	-0.03
SERVICE	0.00	-0.13	0.01	-0.13	-0.03
GOVMNT	0.02	-0.14	-0.22	-0.12	0.21
FOREIGN	-0.30	-0.01	-0.33	-0.01	-0.15
TOTAL	0.00	0.00	0.00	0.00	0.00

3.4.2 Regression Analysis Results

Once we have appropriate measures of the different types of constraints related to the cost and access to bank credit, we can estimate the main determinants of these constraints. I run OLS (robust) regressions with each of the new indicators as dependent variables and the resuts are shown in Table 15. In the same table I also show the results for the Ordinal Probit Estimations of the same model using the original variables HINT and LTLOAN as dependent variables.

The models using the new indicators show a better statistical performance than the ones using the original variables from the survey, if we observe the much higher R-squares. Furthermore, the models using the new variables show more significant results than the ones using the original variables.

The intensity of credit rationing (RATIONED and HINT_NOBANK) clearly decreases with firms' size in all the specifications, since both the signs of SMALL and MEDIUM are positive and significant, and the coefficient of SMALL is larger than the coefficient of MEDIUM. Conversely, the indicators associated with high costs or short maturities (HINT_BANK and RATIONED_LT) increase with the firms' size, i.e. the signs of SMALL and MEDIUM are negative and significant, and the coefficient of SMALL is lower than the coefficient of MEDIUM.

What this is showing is the fact that if smaller firms are the ones that face the worst access to bank credit they cannot be the ones that complain the most about the cost or the conditions of bank loans, since they are basically unable to use any bank credit. It also confirms that the use of expensive alternative sources of credit is more related to small firms, since they are the ones with less access to formal credit from banks.

Something similar happens with the effect of age on RATIONED and HINT_NOBANK versus HINT_BANK and RATIONED_LT. The effect of the firm's age on the access to credit variables (RATIONED and HINT_NOBANK) is clearly negative, although non-linear. The effect of LN(AGE) is negative and significant for both variables as it can be seen in Figure 3.

Conversely, the effect of the firm's age on HINT_BANK and RATIONED_LT is clearly non-linear and they follow an inverse u-shaped evolution with the age of the firm. Both indicators initially increase with the firm's age, but they start decreasing at a certain point and the evolution is thus hump shaped. The RATIONED_LT indicator increases less and decreases faster afterwards. The evolution of these two indicators can be seen in Figure 4.

Therefore, the estimation results suggest that there is a different evolution of the different financial constraints with the age of the firm. After the start-up, as firms become older they increasingly gain access to bank credit and reduce the use of other expensive sources of credit. But as they obtain more bank credit, there is an initial increase in the perception that the costs and the conditions of their loans are a problem for the firm. Such perception reach a maximum point where it starts decreasing and older and larger firms feel much less constrained than medium-age firms. Although this effect could be also seen in the regressions with the original variable HINT, it was not studied or reported by Beck et al (2006) or by any other study that has used the WBES information. A hump shaped evolution of the cost of credit has been reported by Kim, Kristiansen and Vale (2008), but I am not aware of any study that has reported this effect on the perception of paying too high interest rates or any other measure of financial constraints.

Table 15 OLS and Ordinal Probit Regressions with Robust Standard Errors

	rationed	rationed_lt	hint_nobank	hint_bank	hint_nocredit	ltloan	hint
	Number of obs = 5261	Number of obs = 5261	Number of obs = 5517	Number of obs = 5517	Number of obs = 5517	Number of obs = 5261	Number of obs = 5517
	F(13,59) = 18.31	F(14,59) = 8.71	F(13,59) = 19.09	F(14,59) = 14.13	F(13,59) = 6.58	Wald chi2(13) = 191.38	Wald chi2(14) = 129.96
	Prob > F = 0.000	Prob > F = 0.000	Prob > F = 0.000	Prob > F = 0.000	Prob > F = 0.000	Prob > chi2 = 0.000	Prob > chi2 = 0.000
	R-squared = 0.1593	R-squared = 0.0593	R-squared = 0.0936	R-squared = 0.0642	R-squared = 0.0521	Pseudo R2 = 0.0385	Pseudo R2 = 0.0283
small	0.2977***	-0.2840***	0.4449***	-0.3155***	-0.0022	0.0424	-0.0444
	(0.000)	(0.000)	(0.000)	(0.000)	(0.972)	(0.463)	(0.573)
medium	0.1071**	-0.1407***	0.2064***	-0.1568***	0.0698	-0.0043	0.0157
	(0.025)	(0.000)	(0.000)	(0.002)	(0.116)	(0.931)	(0.796)
age		-0.0124***		-0.0119***			-0.0109***
		(0.000)		(0.002)			(0.002)
sqroot_age		0.1155***		0.1265***			0.1063***
		(0.000)		(0.001)			(0.006)
Ln(age)	-0.0815***		-0.0561***		0.0045	-0.0625*	
	(0.001)		(0.009)		(0.829)	(0.064)	
foreign	-0.1237***	-0.1036**	-0.2072***	-0.1307***	-0.0703	-0.1859***	-0.2524***
	(0.006)	(0.038)	(0.000)	(0.010)	(0.180)	(0.000)	(0.000)
government	-0.0079	-0.2223***	-0.1636***	-0.2204***	0.1230***	-0.2021**	-0.2077***
	(0.903)	(0.001)	(0.006)	(0.000)	(0.003)	(0.011)	(0.003)
manufacture	-0.0745	0.0094	-0.0998**	0.0440	-0.0876*	-0.0683	-0.0953
	(0.123)	(0.858)	(0.027)	(0.393)	(0.060)	(0.204)	(0.124)
service	-0.0809	-0.1655***	-0.1126***	-0.1351***	-0.0615	-0.2206***	-0.2448***
	(0.163)	(0.003)	(0.009)	(0.007)	(0.273)	(0.001)	(0.000)
legal_rights	0.0191	-0.0466***	-0.0214	-0.0576***	0.0218	-0.0197	-0.0474*
	(0.472)	(0.003)	(0.114)	(0.001)	(0.258)	(0.500)	(0.088)
cred_inform	-0.0459*	-0.0236	-0.0033	-0.0269	-0.0435**	-0.0553*	-0.0516*
	(0.063)	(0.196)	(0.845)	(0.187)	(0.019)	(0.065)	(0.094)
pub_registry	0.0016	0.0098***	-0.0002	0.0040	-0.0025	0.0094*	0.0016
	(0.698)	(0.005)	(0.958)	(0.394)	(0.466)	(0.069)	(0.782)
priv_bureau	-0.0040***	0.0022	-0.0034**	0.0021	-0.0016	-0.0017	-0.0008
	(0.009)	(0.187)	(0.015)	(0.335)	(0.290)	(0.464)	(0.794)
concentrat	0.5616***	0.4020*	0.2160	0.0817	0.3044	0.7878***	0.3698
	(0.009)	(0.078)	(0.224)	(0.805)	(0.176)	(0.006)	(0.385)
gdp growth	-0.0474***	-0.0170	-0.0252***	0.0040	-0.0229	-0.0593***	-0.0298
	(0.008)	(0.232)	(0.007)	(0.839)	(0.145)	(0.005)	(0.268)
cons	0.0553	0.1033	0.1165	0.2745	-0.0172		
	(0.748)	(0.566)	(0.328)	(0.352)	(0.925)		

It is also important to notice that the size dummies SMALL and MEDIUM are not significant in the regressions run with the original variables HINT and LTLOAN. This can be due to the correlation of these variables with different types of constraints, a problem which is not controlled for if we do not use the new indicators.

Finally, with respect to the size and age of the firm, we find that these variables do not have a significant effect on the HINT_NOCREDIT indicator, which may suggest that this indicator measures a situation that could affect firms of different sizes and ages. I also find that public firms are the ones that complain more about this type of problem. The firm characteristic that is more significantly related to HINT_NOCREDIT is GOVERNMENT. This could be due to public firms having more access to funds from the government, and that they could rely on them if they think that the cost of bank credit is too high.

Another important result is that banking concentration (CONCENTR) has a significant effect on the RATIONED and RATIONED_LT indexes, but not on the HINT_BANK or the HINT_NOBANK indicators. This implies that banking competition (less concentration) improves firm's access to credit but it does not seem to have a significant effect on the interest rates of bank credit. It also implies that banking concentration worsens the general access of firms to bank credit, although it may somehow increase the access in terms of short term credit.

These results seem to give partial support to the structure-performance hypothesis of banking concentration, since I find a positive relation between the main indicator of access to credit (RATIONED) and bank concentration, but no relation between the main indicator of the costs of bank credit (HINT_BANK) and bank concentration.

However, the positive and significant effect of banking concentration on the RATIONED_LT indicator may also be understood as evidence of a certain positive relation between access to credit and banking concentration, since firms scoring high in the RATIONED_LT indicator are firms that do have access to bank loans, although only to short-term ones. The information-based hypothesis says that in highly concentrated markets banks are more likely to offer credit since they can extract monopoly rents from the firms' information collected through relationship lending. However, the information based hypothesis has never considered the possibility that the higher supply of loans offered by banks in an information-monopoly framework could be related to loans of lower quality, such as loans with shorter maturities.

Another implication of the results is that constraints in terms of high costs of credit are found to be lower in markets where regulation facilitates banks to seize collateral assets after bankruptcy. LEGAL_RIGHT coefficient is negative and significant in the HINT_BANK and RATIONED_LT regression. However, it is not significant in the RATIONED, HINT_NOBANK and HINT_NOCREDIT regressions. This means that stronger creditors' protection seems to decrease restrictions by reducing the costs of credit and improving the conditions (longer maturities), but it does not seem to improve firm's general access to credit. In this case we can notice that this variable is not significant if we use the original variable LTLOAN, and that the significance level increases strongly compared to the HINT regression.

CRED_INFORM coefficient is negative and significantly related to the RATIONED indicator and negative and significantly to the HINT_NOCREDIT indicator. The effect of credit information on the other indicators is not significant. These results support the idea that firms in markets with more credit information sharing have more access to credit (less credit rationing) and that information sharing has an effect of increasing the demand for credit of those firms that perceive it as too expensive.

With respect to credit information sharing another interesting result is the different effects of the coverage of private bureaus versus public registries agencies. A higher coverage of private bureaus is found to improve firms' access to credit, since it is negative and significantly correlated to RATIONED and HINT_NOBANK. This result is important since it cannot be found using the original variables in the survey.

Conversely, a higher coverage of public registry agencies is found to increase restrictions in terms of low access to long term credit, since PUB_REGSTR and RATIONED_LT have a positive and significant relation. Similar to the interpretation of the positive effect of banking concentration on the same indicator, the positive relation of RATIONED_LT and PUB_REGSTR could provide evidence that although a higher coverage of public registry agencies does not improve the general access to credit, it does have a certain effect on increasing the access to short-term credit. This difference could be important for policy makers trying to foster financial development by increasing the coverage of both private and public bureaus.

Finally, the results show that that access to credit decreases with worst macroeconomic conditions, but that restrictions in terms of high costs of bank credit do not seem to be affected. GROWTH coefficient is negative and significant in RATIONED and HINT_NOBANK regressions; however, it is not significant in the rest of regressions. This means that under macroeconomic recessions firms have more problems to access credit markets, but the price of credit does not seem to be affected.

Figure 3 Effect of firm's age on RATIONED and HINT_NOBANK indexes according to estimation results

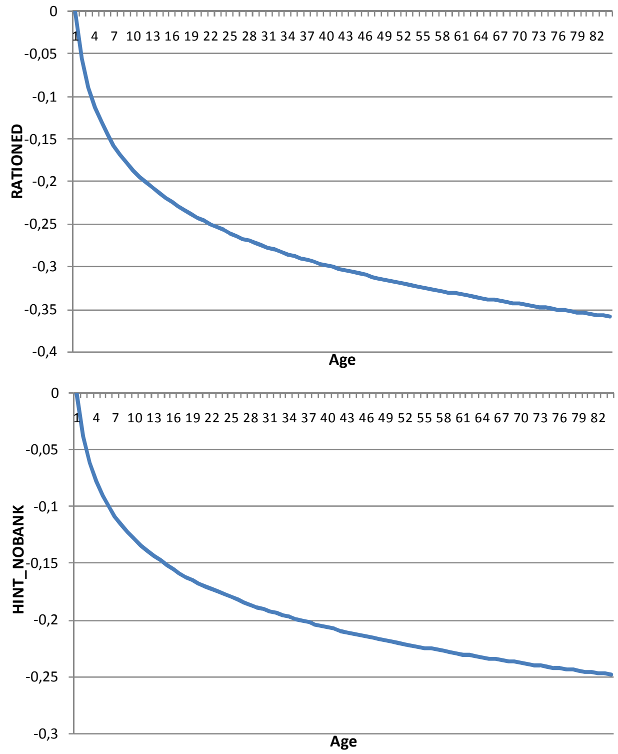
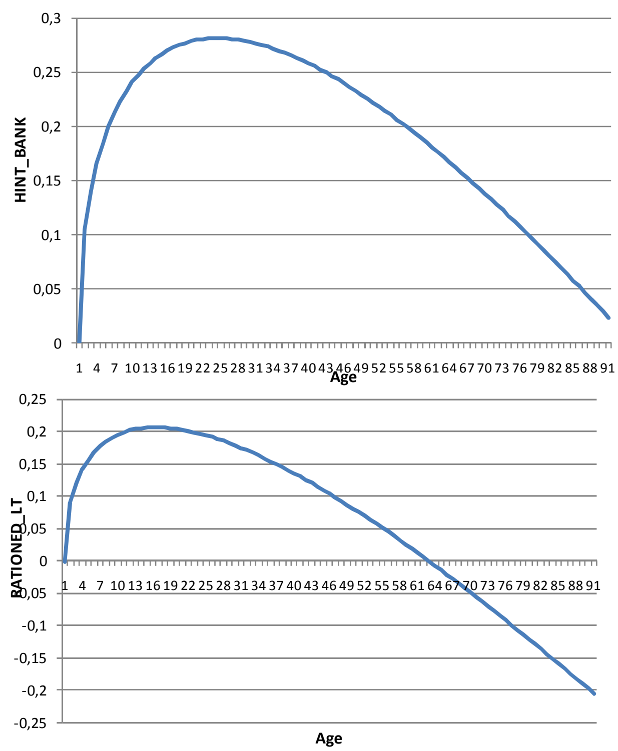


Figure 4 Effect of firm's age on HINT_BANK and RATIONED_LT indexes according to estimation results



3.4.3 Comparison to Previous Studies

As a way to further compare the added value of the methodology and the estimated results to previous studies that have used the same information, I run the same ordinal probit regressions as in Beck et al (2006) using the variables LTLOAN and HINT as dependent variables and using only firm characteristics as independent variables, using country dummy variables to control for common country characteristics. The original results of Beck et al (2006) are shown in Figure 5. Columns "hint(1)" and "ltloan(1)" perform the same regressions as in Beck et al (2006)[4]. The comparison of the results in Table 15 versus Table 16 basically highlights the deficiency in the original WBES survey: some questions in the survey were not very well designed because they could have ambiguous interpretations. If this problem is not correctly controlled for, it can lead to misleading results and to an erroneous analysis if we run an empirical exercise based on those original variables.

For instance, the results about the relation between the variable HINT the firm' size found by Beck at al (2006) does not hold if we control for the use of different sources of credit. In column "hint(2)" I include as independent variables the variables INFORMAL and FN_SCCR (supplier credit)[5]. Controlling for these variables changes the coefficients and the significance of the size dummy variables and the age of the firm. Similarly, in column "hint(3)" I restrict the sample to firms that do not use informal or supplier credit and in column "hint(4)" I only use firms that use bank credit (BANK=1). In both cases the estimated coefficients of the size dummies and the age of the firm are clearly sensitive to these restrictions. The approach of restricting the sample is, however, not correct because most of the firms use different sources of credit at the same time, and because the use of the different types of financing sources it is clearly related to the age and size of the firm.

Similarly, the estimated relationships between the size and age of the firm and the variable LTLOAN do not hold if we control for the use of bank credit. In column "ltloan(2)" I include as explicative variable the dummy variable BANK. In this case, the coefficients of SMALL and MEDIUM increase, but the variable LTLOAN appears to be positive and significantly correlated with the use of bank credit, something difficult to interpret without the CATPCA and its related analysis. Alternatively, in columns "ltloan(3)" and "ltloan(4)" I restrict the sample to firms that use bank credit (BANK=1) in the first case, and firms that do not use bank credit (BANK=0) in the second case. Although the results seem to be reasonable in the case of "ltloan(4)" because small and medium firms are then more constrained than in "ltloan(1)", the opposite happens in the case of the age of the firm that goes from being negative and highly significant to non significant. In column "ltloan(3)" the effect of the size and age of the firm become non significant.

The previous analysis shows that, as explained before, the empirical results in Beck et al (2006) may lead to wrong interpretations of the information in the survey and about the firm

[4] The differences in the results shown in these two columns versus Beck et al (2006) are only due to the different samples used in the two papers.
[5] It is clear that this variable is correlated to the error of the regression and this makes this approach not desirable.

specific determinants of different types of financial constraints. First of all, according to their results, smaller firms are the ones that complain the most about high interest rates. Conversely, according to this chapter results, this is only true if we talk about interest rates from informal lenders or supplier creditors, but not with respect to interest rates from banks. In fact, the complaints about interest rates from bank credit seem to be stronger from larger firms. This does not necessarily means that larger firms pay higher interest rates, but that they are the ones that find banks' interest rates more problematic, meanwhile high banks' interest rates are actually less problematic for the smaller and younger firms. What this chapter shows is that the most important financial problems for the smallest and youngest firms is that they are not able to obtain any bank credit whatsoever and that therefore, they have to pay very high interest rates to informal and supplier creditors.

A similar conclusion follows from the analysis of the access to long term loans. The most important problem for the smallest and youngest firms is not that they do not have access to long term loans, but again, that they do not have access to bank credit at all. Conversely, the problem of receiving only short term lending is a problem more related to larger firms and medium age ones.

Table 16 Ordinal Probit regressions with the original variables and country dummy variables as in Beck et al (2006)

Ordered probit	Number of obs= 5522	Number of obs= 5522	Number of obs= 3435	Number of obs= 2369	Number of obs= 5266	Number of obs= 5266	Number of obs= 2312	Number of obs= 2954
	LR chi2(64) = 1109.03	LR chi2(66) = 1141.18	LR chi2(63) = 695.77	LR chi2(63) = 594.85	LR chi2(64) = 1200.10	LR chi2(65) = 1209.81	LR chi2(63) = 599.39	LR chi2(64) = 704.32
	Prob > chi2 = 0.0000	Prob > chi2 = 0.0000	Prob > chi2 = 0.0000	Prob > chi2 = 0.0000	Prob > chi2 = 0.0000	Prob > chi2 = 0.0000	Prob > chi2 = 0.0000	Prob > chi2 = 0.0000
	Pseudo R2 = 0.0910	Pseudo R2 = 0.0937	Pseudo R2 = 0.0886	Pseudo R2 = 0.1164	Pseudo R2 = 0.0877	Pseudo R2 = 0.0885	Pseudo R2 = 0.0970	Pseudo R2 = 0.0946
	hint (1)	**hint (2)**	**hint (3)**	**hint (4)**	**ltloan (1)**	**ltloan (2)**	**ltloan (3)**	**ltloan (4)**
small	0.0930*	0.0623	0.0684	0.0724	0.1816***	0.2028***	0.0891	0.3015***
	(0.097)	(0.269)	(0.322)	(0.370)	(0.001)	(0.000)	(0.246)	(0.000)
medium	0.1130**	0.1010*	0.0783	0.1687**	0.0679	0.0748	0.0263	0.1417*
	(0.026)	(0.057)	(0.202)	(0.015)	(0.167)	(0.129)	(0.690)	(0.063)
Age	-0.0006	-0.0005	-0.0010	-0.0004	-0.0019**	-0.0018**	-0.0028**	-0.0008
	(0.396)	(0.494)	(0.221)	(0.725)	(0.012)	(0.015)	(0.015)	(0.428)
foreign	-0.2488***	-0.2409***	-0.2412***	-0.2159***	-0.1981***	-0.1958***	-0.2400***	-0.1693**
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.014)
govmnt	-0.1061*	-0.0866	-0.1402**	0.0038	-0.1305**	-0.1272**	-0.1228	-0.1399*
	(0.055)	(0.119)	(0.030)	(0.966)	(0.017)	(0.020)	(0.157)	(0.054)
manufact	-0.0433	-0.0375	-0.0157	-0.0261	-0.0466	-0.0496	-0.1127*	0.0228
	(0.356)	(0.424)	(0.793)	(0.716)	(0.299)	(0.268)	(0.091)	(0.714)
service	-0.2143***	-0.2093***	-0.2429***	-0.1369*	-0.1799***	-0.1756***	-0.1350**	-0.1928***
	(0.000)	(0.000)	(0.000)	(0.053)	(0.000)	(0.000)	(0.043)	(0.001)
bank						0.1104***		
						(0.002)		
informal		0.0036***						
		(0.000)						
fn_sccr		0.0040***						
		(0.000)						

Figure 5 Estimated results in Beck et al (2006)

	(1) High interest rates	(2) Collateral requirements	(3) Access to long-term loans
Age	−0.006	−0.062	−0.042
	(0.30)	(3.21)***	(2.02)**
Small	0.133	0.246	0.150
	(2.03)**	(4.19)***	(2.99)***
Medium	0.081	0.149	0.027
	(1.46)	(2.88)***	(0.58)
Listed	0.095	−0.091	−0.035
	(1.69)*	(1.74)*	(0.47)
Group	−0.007	0.038	−0.141
	(0.10)	(0.61)	(1.62)
Foreign	−0.260	−0.252	−0.242
	(6.61)***	(5.76)***	(4.98)***
Government	0.103	−0.048	−0.020
	(1.60)	(0.75)	(0.33)
Multinational	−0.068	0.025	−0.005
	(1.61)	(0.64)	(0.09)
Pseudo R^2	0.103	0.052	0.100
Observations	6163	5833	5507

3.5 Conclusions

This chapter investigates the main determinants of different types of financial constraints related to low access to credit (credit rationing and the use of expensive sources of credit) and restrictions in terms of adverse conditions in bank loans (high interest rates and short maturities). The study of the determinants of financial constraints has always been limited by the lack of generally accepted measures of financial frictions. Recent studies such as Beck et al (2004) and Beck et al (2006) use survey-based information to explore the determinants of firms' financing obstacles. However, the analysis of some variables/questions used in these studies shows that they could present deficiencies in these original variables that weaken their conclusions.

More specifically, the variables from the survey that aim to measure how problematic interest rates and the access to credit are could have very ambiguous interpretations, because firms facing different types of problems can reply in the same way to different questions. If this problem is not correctly controlled for, it can lead to misleading results and to erroneous conclusions about the determinants of specific types of financial constraints. In order to correct this problem, I construct new indicators based on the two variables mentioned before, together with other variables included in the survey that help clarifying which specifc problem each firm is facing. Thus, the new constructed variables are better indicators of each specific type of financial problem, such as how problematic interest rates from banks are, how problematic interest rates from informal lenders or supplier credit are, how problematic not having any access to credit is, and how problematic having access only to short-term financing is.

For instance, the results in Beck et al (2006) suggest that smaller firms are the ones that complain the most about high interest rates. Conversely, according to this chapter, this is only true if we talk about interest rates from informal lenders or supplier creditors, but not with respect to interest rates from banks. In fact, I find that the stronger complaints about interest rates from banks' credit come from larger and medium age firms. The new indicator that measures this problem actually displays an inverted u-shaped relationship with the age and size of the firm. This does not necessarily means that larger firms pay higher interest rates than the smaller ones, but that larger firms are the ones that find banks' interest rates more problematic, meanwhile high interest rates from banks are actually less problematic the smaller and younger a firm is. What this chapter shows is that the most important financial problem for the smallest and youngest firms is that they are not able to obtain any bank credit whatsoever and that therefore, they have to pay very high interest rates to informal and supplier creditors.

A similar conclusion follows from the analysis of the access to long-term loans. In Beck et al (2006) this variable is negatively related to the age and size of the firms. This result does not hold with the new indicators. According to this chapter's results, general access to bank credit monotonically decreases with the age and size of the firm. However, the specific problem of receiving only short-term lending is a problem more related to larger firms and medium age ones. The new indicator that measures this problem actually displays an inverted u-shaped relationship with the age and size of the firm. Therefore, the results highlight again that the most important problem for the smallest and youngest firms is not that they do not have access to long-term loans, but again, that they do not have access to bank loans at all.

The chapter also explores how country characteristics relate to the specific financing problems described before. Although, to the best of my knowledge, there are no other studies that investigate how country characteristics relate to these specific financial problems, the results are interrelated to those that study the relation between banking concentration and the access and cost of credit. Regarding firms' access to credit and banking concentration, the results are consistent with the structure performance hypothesis (e.g. Pagano 1993 or Guzman 2000) that states that more concentration reduces access to credit. However, in our results, concentration does not seem to have a significant effect on the constraints related to the cost of bank credit. A novel result of this chapter with respect to the previous literature is that although banking concentration worsens the general access of firms to bank credit, we find some evidence that it may increase the access in terms of short term credit.

Regarding the other determinants of the specific financial problems evaluated here, it is also found that the perception of facing unfavourable conditions of bank credit is smaller in markets where regulation facilitates banks to seize collateral assets after bankruptcy. However, I did not find a significant effect on the indicators related to credit rationing. This means that stronger creditor protection seems to improve financial restrictions by reducing the costs and improving the conditions of bank loans, but it does not seem to improve firms' access to bank credit.

The results also support the idea that firms in markets with more credit information sharing have more access to credit (less credit rationing), although it does not have an effect on the restrictions in the restrictions related to adverse conditions of bank loans. Additionally, more credit information sharing is found to increase the demand of credit from firms that think that credit is too expensive.

With respect to credit information sharing, I also find different effects of the coverage of private bureaus versus public registries agencies. A higher coverage of private bureaus is found to improve firms' access to credit, but on the contrary, a higher coverage of public registry agencies is found to increase restrictions in terms of low access to long term credit, which can imply that a higher coverage of public registry agencies only helps improve the access to credit with short-term or adverse conditions.

Finally, I find that credit rationing worsens with worst macroeconomic conditions. However, I do not find a significant effect on the indicators of high costs of bank credit or restrictions in the maturity. This means that access to cred.it seems to be pro-cyclical, but the perception of paying too high costs for credit is not affected by the economic cycle.

CHAPTER 4

4. NON-MONOTONIC RELATIONSHIP BETWEEN INVESTMENT AND INTERNAL FUNDS IN A MODEL OF INVESTMENT WITH FINANCIAL CONSTRAINTS AND LEARNING

In this chapter I develop a new framework for analyzing the possibility of asymmetric learning between a bank and a firm and its effect on firm's investment decision. I use this framework to investigate theoretically and empirically the relationship between firms' investment and internal funds in the presence of limited information, learning and bankruptcy costs. The chapter provides new arguments to support the theory of a U-shaped relation between investment and internal funds, and it does it in a more intuitive and general way than in previous theoretical models. On the empirical side, this chapter presents new evidence of the non-monotonic relation between investment and internal funds. According to the predictions of the theoretical model, I find that this non-monotonicity is stronger for younger firms and that the negative relation does not depend on the sign of internal funds.

4.1 Introduction

In this chapter I develop a new framework for analyzing the possibility of asymmetric learning between a bank and a firm and its effect on a firm's investment decisions. I apply this framework to investigate theoretically and empirically the relationship between firms' investment and internal funds in the presence of limited information, learning and bankruptcy costs.

The recent literature in corporate finance has not yet clarified how and/or why financial constraints have a real effect on investment. Even though many authors have shown that investment cash flow sensitivities are a very noisy measure of the intensity of financial constraints, these are still considered useful in evaluating the real effect of financing frictions, because of the lack of more precise measures. Regarding this, recent empirical studies suggest that a firm's investment level is not a monotonically increasing function of internal funds, but instead that is a u-shaped function, because investment actually decreases with internal finance for the lowest levels of wealth.

To the best of my knowledge there is only one theoretical model that provides a possible explanation of why investment could be a u-shaped function of the level of internal funds, which is the model of Cleary, Povel and Raith (2007). These authors consider the model of a firm with idiosyncratic risk and with costly bankruptcy. However they derive the U-shape investment function under very restrictive assumptions, since their results only hold if the revenues shocks follow the uniform distribution.[6]

In this chapter I show that the model with asymmetric learning is able to provide a more robust explanation of the U-shape in the investment-internal finance relation. More specifically, this chapter provides the following new results:

i) It shows that a U-shape investment function arises naturally in the presence of asymmetric learning of banks and firms about the firms' riskiness, even in the absence of bankruptcy costs.
ii) It extends Cleary *et al* (2007) model in order to clarify under what conditions bankruptcy costs are indeed sufficient for generating the U-shape function under more general distributional assumptions.
iii) It calibrates with realistic data both the model with asymmetric learning and the model with bankruptcy costs, to show that the U-shape function is more likely to arise if asymmetric learning is present.
iv) It estimates the U-shape function empirically with firm level data, to provide further evidence that asymmetric learning is the most likely explanation of it.

The new investment model proposed in this chapter shares the same three basic assumptions described by Cleary *et al* (2007): First, external funds are more expensive than internal funds

[6] Moreover, the u-shape result in Cleary *et al* (2007) requires the unrealistic feature of the model that when internal funds are negative capital could become cheaper the more the firm invests, because the probability of default decreases with a higher investment. This follows from the fact that when internal funds are negative, the firm will default for sure if it does not invest at all or if it invests too little. Therefore, the interest rate decreases the more the firm invests, even though it has to demand larger amounts of debt.

because of the presence of an asymmetry of information or other capital market inefficiency. Second, the cost of raising external funds is endogenously determined by the creditor's requirement of receiving the same expected return when debt is risky than when debt is fully secured and risk-free. Third, the firm can choose any desired level of capital instead of simply choosing between investing or not in an indivisible project. Besides these three main assumptions, the two models also have in common that liquidation of the firm is inefficient, since the assets are worth more to the owner than to the creditor.

The main differences with Cleary *et al* (2007) are as follows: first, I assume that the bankruptcy cost is proportional to the value of investment and that, at the same time, firm revenues follow a log-normal distribution. Both assumptions are at the same time more realistic than Clearly's assumptions and are sufficient to imply that bankruptcy costs still generate a U-shaped investment function.[7]

Second, in this model I also consider a different type of capital market imperfection which is totally independent of the effect of the bankruptcy costs. I consider the possibility that creditors and firms do not share the same set of information for predicting the firm's probability of default. Therefore both firms and banks have to learn about the firm's profitability and about the probability of debt repayment in the presence of financial imperfections. Learning is a realistic assumption widely used in other economics fields. Firms and banks do learn about the firms' profitability, and yet this feature has not been considered in the financial frictions literature. More specifically, in the model firms and banks use historical information about a firm to learn and predict its future profits and default probability.

This chapter solves this model and shows that investment is non-monotonically related to internal funds in a u-shaped fashion. However, it is quite important to clarify that not all assumptions are needed jointly to generate the u-shape relationship. So by adding all these ingredients I define two possible explanations for the u-shape: the asymmetric learning and the non-monotonicity of the density function with respect to internal funds and its interaction with the expected bankruptcy costs.

I thus find that investment is increasing on the level of internal funds for those firms with a medium-low, medium and medium-high level of internal funds, but that investment could be decreasing on internal funds for those firms with the lowest levels of internal funds (whether they are negative or not). Finally, investment should not change with internal funds for firms with a very high level of it (unconstrained firms).

[7] Importantly, I obtain a u-shaped investment curve in a setting in which, at the optimum, the firm cannot decrease its marginal costs by investing more, but on the contrary, if it invests more than its optimal level it would have to pay a higher interest rate because the probability of default would be higher. Furthermore, I show that in all the different cases analyzed, it is impossible that the probability of default decreases with a higher investment at the optimal level of capital, even if internal funds are negative. In the present model, when internal funds are negative the probability of default could still be decreasing on investment, but only for some levels of capital lower than the optimal one, because at those levels, the revenues from investment are not enough to repay the existing burden of debt represented by a negative financial wealth.

The rationale for the u-shaped relationship in the case that the bank has a shorter history of information than the firm (asymmetric learning) is the following: Any disagreement between a bank and a firm matters most when repayment or default are equally likely, and it matters less when any of them are almost certain.

Intuitively, if the firm itself is not sure if it will default or if it will be able to repay its debt because it predicts that both situations are equally likely, any difference in the creditor's information could make a large difference in its prediction about the likelihood of those events, compared to the firm's beliefs. On the other hand, if either default or repayment are almost certain, the predictions from both sides are going to be quite similar independently of the set of information that they are using.

Therefore, the effect of a disagreement between a creditor and a firm caused by the creditor's higher uncertainty (variance) depends on the level of internal funds: it is more important for intermediate levels of risk (intermediate levels of internal funds) and it is less important for high and low levels of risk (low and high levels of internal funds).

In the model, this translates into the following: When bank and firm share the same information they predict the same probability of default, and therefore, they both agree on the fair cost of debt independently of how risky investment is, i.e. independently of the level of internal funds. However, with asymmetry of information the bank may estimate a lower probability of repayment, or the bank may expect to recover a lower value of revenues in the case the firm defaults, compared to what the firm expects to transfer to the bank in that case. When any of those situations occurs (or both) the bank charges the firm a higher interest rate than what the firm considers fair. Therefore, the marginal cost of investment is higher than in the symmetric case and the firm must decrease its level of capital to increase its marginal revenue and compensate for the higher marginal cost.

The chapter first shows that the asymmetry of information always translates into a higher marginal cost (or a lower marginal product) and thus investment is always smaller in the asymmetric case than in the symmetric one. Furthermore, since the disagreement is lower when internal funds are high or low, and higher for intermediate levels of internal funds, the difference in the marginal cost between the asymmetric and symmetric cases is lower when internal funds are high or low, and higher for intermediate levels. Thus, the difference in investment between the asymmetric and symmetric case is also smaller in the extreme cases and higher (lower investment) in the intermediate ones. Clearly, investment is thus non-monotonically related to internal funds in a u-shaped fashion.

The intuition of the effect of the asymmetry of information can also be illustrated with an example of two gamblers betting on a soccer match. We can imagine two different gamblers trying to predict the probability that team A defeats team B. If we allow them to have different sets of information, their estimated probabilities may be different as well. However, the difference in their estimated probabilities is not independent of the score of the game. For instance, if team A is winning by a 10 goals difference it is obvious that both gamblers will agree that team A is more likely to win the game. Similarly, if team A is losing by a 10 goals difference, again, both gamblers will surely agree that team A will lose the game. On the other hand, we may expect that the difference in their estimated probabilities should be larger when the game is tied. Therefore, we should expect that the disagreement between the two is more important when both teams have the same probability of winning, which occurs when they are tied.

The rationale for the second possible explanation for the u-shaped relationship, the positive bankruptcy costs, is similar to the previous one, although totally independent of whether we assume asymmetry or symmetry of information: The effect of a marginal decrease in internal funds on the predicted probability of default is higher when repayment or default are similarly likely, and lower when any of them are almost certain.

Therefore, the effect of a marginal change in internal funds and in firm's leverage depends on the level of internal funds: it is more important for intermediate levels of risk (intermediate levels of internal funds) and it is less important for high and low levels of risk (low and high levels of internal funds).

In the model, this translates into the following: when internal funds are higher than the optimal level of investment, the firm does not require external funds and the probability of default is zero. If internal funds decrease a little and the firm has to demand some external funds, the default probability increases from zero to a positive level. As internal funds are high and the level of risk is very low, a marginal decrease in internal funds generates only a small increase in the probability of default and in the marginal cost. As the marginal cost of capital increases, the firm has to reduce the level of capital to increase its marginal revenue and to compensate for the increase in the probability of default and the interest rate.

However, the increase in the probability of default becomes larger as internal funds decrease further and the level of risk increases. Therefore, the decrease in investment has to be larger with every change in internal funds to compensate for the incremental increase in the marginal cost. Such incremental increase in the probability of default reaches a maximum when the level of risk reaches the mode of the distribution (the most frequent value).

If internal funds decrease enough to take the level of risk above the mode of the distribution, any further decrease in internal funds generates a smaller increase in the probability of default than before, and therefore, the decrease in investment becomes smaller with every change in internal funds. At this point, the level of capital is very low, and the marginal return per unit of capital is, consequently, very high. When internal funds decrease further enough and the increase in the probability of default per unit of capital is really small, the high marginal return per unit of capital is large enough to compensate for the increase in the marginal cost per unit of capital, and therefore, the firm prefers to increase investment.

It is important to notice that the results imply that the probability of default and the interest rate always increases with an increase in leverage, but the increase is smaller when risk is very high or very low. Furthermore, I show that this result does not depend on the whether internal funds are negative or positive.

Again, the intuition of the non-monotonicity of the expected bankruptcy cost can also be illustrated with an example of a soccer match. We can compare the effect of a marginal change of internal funds for a firm's estimated probability of default to the effect of an additional goal for a gambler's estimated probability that a team "A" will win against team "B". In general, if team A scores one goal, this always increases the gambler's estimated probability that team A will win the game. However, the effect of one goal on the estimated probability is not independent of the score of the game at the moment that it occurs. For instance, if team A is already winning by a 10 goals difference, the effect of the 11th goal will not increase too much the gambler's belief about team A winning because the probability is already close to one. Similarly, if team A is losing by a 10 goals difference, the effect of a 11th goal from team B does not change too much the estimated probability of team A winning or losing. On the contrary, if the game is tied, one goal from team A largely increases the estimated probability of team A winning the game.

Applying this idea to the bank's estimated probability of a firm's default, it is clear that the effect of a marginal change in internal funds should be larger when default and repayment are similarly likely than when any of them are almost certain. Since the marginal cost of external funds should depend on the probability of default and thus on any change in internal funds, the marginal cost of investment may change with internal funds in a non-monotonous way.

To test the theoretical predictions of the model, I first simulate an artificial sample of firms based on the introduction of realistic parameters to the theoretical model, in order to compare which one of the two alternative explanations for the u-shape investment curve is more likely to drive the empirical evidence. The results of the simulation imply that the asymmetric learning story is the one that could explain more clearly the existence of a u-shaped investment curve.

Finally, in the third part of the chapter I estimate a regression model of fixed capital investment using an unbalanced panel containing 257,566 firm-year observations of SABI data (SABI accounts for Iberia Balance Sheet System Analysis).

I conduct a similar empirical exercise to previous studies and perform three kinds of tests. First, I regress the log of fixed capital on the lagged value of financial wealth and on other variables measuring the level of internal funds, while controlling for investment opportunities. For accounting for the non-linear relationship between investment and internal funds, I conduct spline regressions of investment on financial wealth. Thus, I estimate investment as a piecewise linear, continuous function of wealth by splitting the data into different sixtiles and other ranges of internal funds. In all regressions, predicted investment is first decreasing in wealth for the lowest levels of wealth and increasing afterwards.

Second, as an alternative way to detect nonlinearities in the data I estimate a polynomial specification that includes the squared value of lagged wealth. Consistent with the idea of a non-monotonic relationship, I find that both coefficients for the wealth (linear and squared) are significant, that the coefficient for the square term is positive and that including the polynomial term improves the explanatory power of the regression.

Finally, similar to many other studies in the investment literature, I run split-sample regressions. Specifically, I regress investment on internal funds separately for observations with the lowest values of internal funds versus observations with the highest values of internal funds. Consistent with the predictions, I obtain a positive coefficient for the high wealth group and a negative coefficient for the low wealth one.

The importance of the chapter results is mainly theoretical, but also empirical. Not only provides additional arguments to support the u-shaped curve theory of investment and internal funds, but it also does it in a more intuitive and general way than in Cleary *et al* (2007), which is the only other model with a similar theoretical prediction. Additionally, it contributes to the general theory by providing a completely new framework for analyzing the possibility of asymmetric learning and the estimation of probabilities with different information sets.

On the empirical side, this chapter confirms the presence of a negative relation between internal funds and investment for firms with the lowest levels of internal funds. An important contribution is that I find this evidence using a sample of firms which do not have access to stock markets and therefore depend more on external funds from financial intermediaries, which made them more sensible to the kind of financial imperfections studied in the theoretical model.

According to the predictions of the theoretical model, the empirical evidence confirms that young firms, which are the ones more likely to face stronger asymmetries of information and higher uncertainty in their distribution of profits, show a clearer u-shaped relation between investment and internal funds. Therefore, both the simulation results and the empirical evidence point in the direction of the asymmetric learning explanation. Importantly, the spline regressions show that not all the firms with negative values of internal funds show a negative relation between investment and the different measures of internal funds, a result that further supports the theory presented herein.

4.2 Theoretical Model

The theoretical framework is based on a model of firm investment, with two times $t = 0,1$. At $t = 0$, the firm has to decide on the level of capital k_1 for next period ($t = 1$). The firm has a standard concave production technology $f(k_t, \theta_t) = e^{\theta_t} k_t^{\beta} = \Theta_t k_t^{\beta}$.

The parameter θ determines the true profitability of the firm, meanwhile the observed profitability can vary because of the randomness in the variable θ_t (which may come from the randomness in productivity and sales). The firm does not know the true value of θ and therefore it needs to learn about it based on the observation of previous realizations of θ_t.

θ_t is normally distributed, i.e. $\theta_t \sim N(\theta - \sigma^2/2, \sigma^2)$, therefore $e^{\theta_t} = \Theta_t$ is log-normally distributed: $e^{\theta_t} = \Theta_t \sim \log N\left(\exp(\theta), (\exp(\sigma^2) - 1)\exp(2\theta)\right)$.

After estimating θ, the firm can predict the value of its profits at time $t = 1$ and it can use this prediction to decide about k_1. I assume for simplicity that the firm knows the true value of σ^2 so it only needs to learn about θ. If the firm did not know about σ^2, there would be a similar process of learning about this parameter, and the posterior probability distribution for the prediction would be of a different type, but the main results presented here would not change.

The key issue of the model is that given that some parameters have to be learned for the prediction of the firm's income shock, then the variance of the prediction depends on the information set used to estimate the parameters, and more specifically, the variance depends negatively on the number of observations that are used in the estimation.

In the case of a new firm the variance decreases with higher experience of the entrepreneur, and for established firms the variance decreases with the age of the firm.

If for instance, a creditor (a bank) needs to estimate the same parameters and has a different information set, then the probability distribution of its prediction will depend on the information set the creditor possesses about the firm's profitability. In general, the bank should have less or equal information than the firm about the firm's profitability.

Given that the information set of firms and banks may be different, then the variance of the prediction of the bank may also be different. Therefore, the distribution function of the prediction may differ between bank and firm and the estimated probability of any given point of the prediction's distribution (such as the default point) may also be different.

The Bayesian learning framework shows that we can summarize any effect of the learning process (or accumulation of a history of information) in the evolution of the posterior variance or the variance of prediction. As firms acquire more information about themselves they can learn that they are better, the same, or worst than what their prior information says. Moreover, with more information uncertainty about the true firm quality is reduced and therefore, the prediction of future profits is more precise. Actually, after learning which the true firm quality is, the only uncertainty about the future comes from the simple randomness in profits (sales or productivity) but the uncertainty about how good a new firm is, disappears.

A basic assumption of the model is that there is perfect information sharing among banks and that there is perfect competition in the credit market. Therefore, I am not considering the possible effects of banking competition in determining an optimal contract for the firm when there is an asymmetry of information. I am also not considering the possible lock-in or hold-up problems arising from the learning process.

Of course, in a more complete framework, we have to take into account the competitive conditions of the credit market and the possible lock-in or "monopoly" power arising from the acquisition of information when information sharing is not perfect. In those cases, a given bank may profit from a contract that makes firms reveal all the extra information they have about themselves in the case of asymmetry of information. This possibility should be proved and explored, but I do not do it here. In this chapter I am only interested in the pure effect of learning. That is, in having a longer versus having a shorter history of information, and in the effect that this has on the relative perception of risk.

4.2.1 Firm's Learning Process

The firm can observe θ_t when it observes the realizations of its revenue function $f(k_t, \theta_t) = e^{\theta_t} k_t^{\beta}$, given that I assume the firm knows its own technology β. Thus, if at $t = 0$ the firm has lived for T periods, it has also observed T realizations of θ_t. Therefore, the firm's information set about θ_t is denoted by $\Omega_F = \{\theta_0, ..., \theta_{-T}\}$ where T is equal to the age of the firm.

Using the standard formulas for normal prior-normal information distributions (see Zellner (1971), O'Hara (1995), Bernardo (1994)), the posterior probability distribution function for θ will be normally distributed with mean

(1) $$E_0(\theta \mid \Omega_F) = \frac{(\hat{\theta})(\sigma^2 / T)^{-1} + (\theta_{0F} - \sigma_{0F}^2 / 2)(\sigma_{0F}^2)^{-1}}{(\sigma^2 / T)^{-1} + (\sigma_{0F}^2)^{-1}}$$

and variance given by

(2) $$\text{var}(\theta \mid \Omega_F) = \left(\frac{1}{(\sigma^2 / T)^{-1} + (\sigma_{0F}^2)^{-1}} \right)$$

where $\hat{\theta} - \sigma^2$ is the estimated sample mean of the observed θ_t, and $\theta_{0F}, \sigma_{0F}^2$ represent the mean and the variance of the prior probability distribution function for θ.

In Appendix A. 2 I derive the prior distribution of firms. The prior distribution basically depends on the population average ability and the industry average profitability.

In equation (1) I am imposing that the prior distribution for θ should have a mean preserving distribution in the sense that changes in the prior variance should not change the expected value of the final log-normal distribution, i.e. I am assuming that the prior distribution for θ is $N(\theta_{0F} - \sigma_{0F}^2/2, \sigma_{0F}^2)$.

Based on the observation of previous realizations of θ_t and on his prior beliefs about the distribution of θ, an entrepreneur can predict the future value of θ_t. Denoting the future (unobserved) value of θ_t at $t=1$ as θ_1 and its predicted value as $\tilde{\theta}_{1F}$, we have that $\tilde{\theta}_{1F}$ is given by:

(3)
$$\tilde{\theta}_{1F} = E_0(\theta_1 - \sigma^2/2 \mid \Omega_F) = E(\theta + \varepsilon_1 - \sigma^2/2 \mid \Omega_F)$$
$$= \frac{(\hat{\theta})(\sigma^2/T)^{-1} + (\theta_{0F} - \sigma_{0F}^2/2)(\sigma_{0F}^2)^{-1}}{(\sigma^2/T)^{-1} + (\sigma_{0F}^2)^{-1}} - \sigma^2/2$$

and a has a variance σ_F^2 given by:

(4)
$$\sigma_F^2 = \text{var}(\theta_1 - \sigma^2/2 \mid \Omega_F) = \text{var}((\theta + \varepsilon_1 - \sigma^2/2) \mid \Omega^F) =$$
$$= \text{var}(\theta \mid \Omega^F) + \text{var}(\varepsilon_1) + \text{var}(-\sigma^2/2) = \left(\frac{1}{(\sigma^2/T)^{-1} + (\sigma_{0F}^2)^{-1}}\right) + \sigma^2$$

where the normal random error ε_1 corresponds to the error of prediction at t = 1 which is distributed with mean zero and variance σ^2.

Given that $\theta_1 \sim N(\tilde{\theta}_{1F}, \sigma_F^2)$, we have that $\Theta_1 = e^{\theta_1}$ is log-normally distributed with an expected value given by:

(5)
$$\tilde{\Theta}_{1F} = E_0(\Theta_1 \mid \Omega_F) = \exp\left(\tilde{\theta}_{1F} + \sigma_F^2/2\right)$$

And a variance given by:

(6)
$$\text{var}(\Theta_1 \mid \Omega_F) = \left[\exp(\sigma_F^2) - 1\right]\exp(2\tilde{\theta}_{1F} + \sigma_F^2)$$

If we replace $\tilde{\theta}_{1F}$ and σ_F^2 given by equations (3) and (4) into equations (5) and (6) it can be shown that $\Theta_1 = e^{\theta_1}$ is log-normally distributed with an expected value equal to:

(7)
$$\tilde{\Theta}_{1F} = E_0(\Theta_1 \mid \Omega_F) = \exp\left(\theta_F\right)$$

And a variance equal to:

(8)
$$\text{var}(\Theta_1 \mid \Omega_F) = \left[\exp(\sigma_F^2) - 1\right]\exp(2\theta_F)$$

Where $\theta_F = \dfrac{(\hat{\theta})(\sigma^2/T)^{-1} + (\theta_{0F})(\sigma_{0F}^2)^{-1}}{(\sigma^2/T)^{-1} + (\sigma_{0F}^2)^{-1}}$.

This is also equivalent to say that θ_1 is normally distributed as $\theta_1 \sim N(\theta_F - \sigma_F^2/2, \sigma_F^2)$.

In Figure 6 and Figure 7 we can see an example of the learning process. In this example, there are three true parameters (three firms), θ could be equal to 0, 0.1 and -0.1. The known σ^2 is equal to 0.1. The initial variance is the same for the three firms and is equal to 0.2. Therefore, the mean of the normal prior distribution for θ_1 is -0.1 for the three firms.

The evolution of the normal mean is shown in the two solid lines and the long dashed line, and the evolution of the variance is represented by the dotted line and is the same for the three firms. We can see that the mean of the normal distribution for the firm with $\theta = 0$ starts at -0.1 and as the firm learns it approaches the true mean of the normal distribution which is $\theta - \sigma^2/2 = 0 - 0.05 =- 0.05$. Meanwhile the variance of prediction starts at 0.2 and through learning it goes down until approaching the true value 0.1

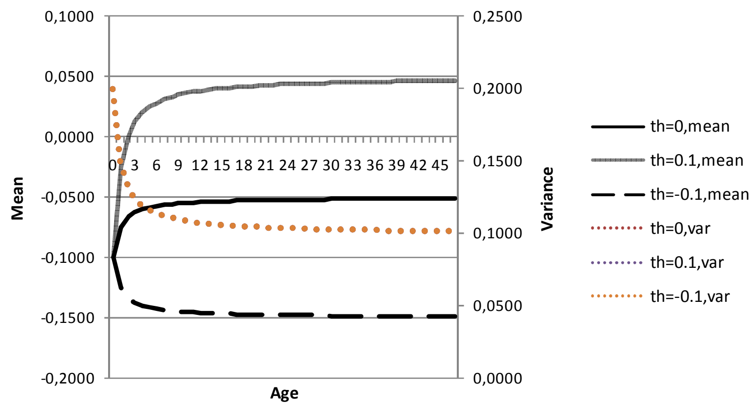


Figure 6. Learning Process, normal distribution $\theta_1 \sim N(\theta_F - \sigma_F^2/2, \sigma_F^2)$. Prior mean is equal to 0, prior variance and true variance are equal to 0.1, and three different firms with $\theta = 0$, $\theta = 0.1$ and $\theta = -0.1$

Similarly, the evolution of the mean of the log-normal distribution for all three firms starts at exp(0) = 1. When there is nothing to learn, which is the case of $\theta = 0$, the mean of the log-normal distribution remains at exp(0) = 1. However the variance of prediction of the log-normal does decrease because of the learning process.

For the firm with $\theta = 0.1$ the mean of the log-normal increases with learning until approaching its true value, exp(0.1). The same for the firm with $\theta = -0.1$, the mean of the log-normal decreases until it approaches its true value exp(-0.1). The variance for all the firms decreases as the firm learns its true type.

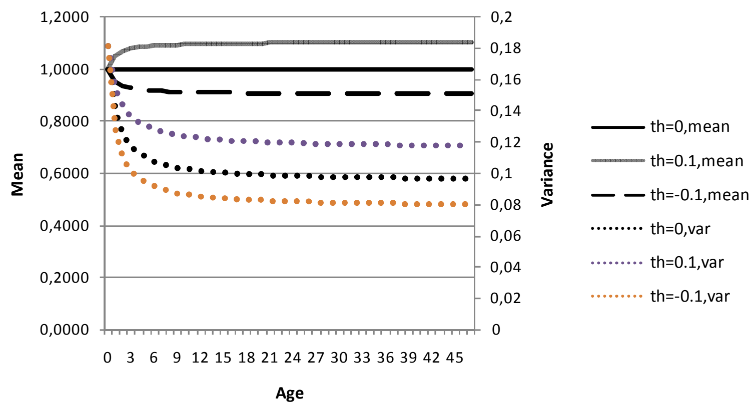


Figure 7. Learning Process, Log-normal distribution, $\tilde{\Theta}_{1F} = E_0(\Theta_1 \mid \Omega_F) = \exp(\theta_F)$. Prior mean is equal to 0, prior variance and true variance are equal to 0.1, and three different firms with $\theta = 0$, $\theta = 0.1$ and $\theta = -0.1$

4.2.2 Investment Decision and First-Best Case

I can express the value of firm's assets at the end of a given period through a parameter q, with $q = (1 - \delta - \varphi)$, where δ is the depreciation rate of capital and φ is the intangible part of investment assets. The intangible part of assets is lost when the firm exits the market or if the firm wants to sell its assets.

The net amount of funds available to invest at $t = 0$ is equal to W_0, where W_0 represents the initial value of equity. $R = (1 + r)$ is the risk-free rate.

Investment level is given by k_1. If the firm has enough resources to finance investment, i.e. if $W_0 > k_1$, the firm does not need to ask for external finance and firm only decides on the level of k_1 that maximizes the expected value of equity at $t = 0$ which is equal to:

(9) $$\max_{k_1} E_0 V_0 = W_0 - k_1 + E_0\left[\frac{e^{\theta_1} k_1^{\beta} + q k_1}{R}\right] = W_0 - k_1 + \frac{\tilde{\Theta}_{1F} k_1^{\beta} + q k_1}{R}$$

The first order condition for k_1 is given by:

(10) $$\beta \tilde{\Theta}_{1F} k_1^{\beta - 1} = R - q$$

4.2.3 Need for External Finance, Default and Bankruptcy Costs

Now I assume that some amount of external financing B_1 is needed to finance the next period capital level k_1, because $W_0 < k_1$. In this case the firm would have to repay an amount RB_1 at $t = 1$, where $R = (1 + r)$ is the risk-free interest rate. The value of debt would be $B_1 = k_1 - W_0$.

In the case that realized revenues turn out to be much smaller than expected, firms can self-liquidate assets to repay the debt, selling them for a value of qk_1, thus, losing the part of assets that is intangible[8]. Additionally, in the case the firm may default, the bank has the right to liquidate the firm. However, if the firm is liquidated, the net value of the firm after liquidation is reduced because the creditor and the firm have to pay for bankruptcy costs, like audits or legal costs.

The proportion of assets lost due to bankruptcy costs is called τ, where $\tau \in [0,1]$ This is similar to Cooley and Quadrini (2001) but in their model the bankruptcy cost is expressed in absolute terms, meanwhile I am expressing the bankruptcy cost as a proportion of investment. Therefore, the net value of the firm in liquidation (inside collateral) will only be equal to $(1-\tau)qk_1$ similar to Almeida and Campello (2002).

In the case of no default, $RB_1 \leq (1-\tau)qk_1$, the problem of the unconstrained firms would be the following:

(11) $$\max_{K_1} \mathrm{E}_0(V_0) = \frac{1}{R} E_0\left[e^{\theta_1} k_1^{\beta} - (RB_1 - qk_1)\right] = \frac{1}{R}\left[\tilde{\Theta}_{1F} k_1^{\beta} - (RB_1 - qk_1)\right]$$

with $B_1 = k_1 - W_0$,

The first order condition for this problem is the same as in (10).

From now on, I called the expected value of equity defined at equation (11) and the level of capital k_1 defined by equation (10) as the first-best case (FB).

Whenever banks know that the firm may default after a bad realization of θ_1, they know that they would incur in losses. Therefore, debt is risky and banks cannot charge the risk-free interest rate. All the agents in the economy are risk-neutral (firms and creditors).

I assume that there is perfect information sharing among banks. Any information that a bank has about a firm would be immediately known by a competitor bank. I also assume that there is perfect competition among banks, although I only need to assume that there is no monopoly in the banking sector and that information sharing is perfect.

Thus, banks would set loan contract terms according to the following rule: firms can borrow any amount of credit $B_1 \leq k_1 - W_0$ at the risk-free interest rate *R,* whenever $RB_1 \leq (1-\tau)qk_1$. What this means is that creditor will always receive the promised repayment, no matter what the state of nature happens, i.e. whatever the realization of the income shock. If $RB_1 > (1-\tau)qk_1$ firms will be able to borrow some amount of uncollateralized debt, but at a higher interest rate than *R*, which I will denote by R_1.

[8] Tangible assets are mainly accounts receivables, inventories, equipment and real estate (Berger and Udell (1998)).

Given that banks are risk-neutral and that there is perfect information sharing among banks, the rule for the loan's interest rate R_1 that the bank will charge for unsecured debt, is simply determined by a non-arbitrage condition and thus, the repayment that the bank can get in the case for non-risky debt should be equal to the expected value of the repayment in the risky case (R_1).

The expression $RB_1 = (1-\tau)qk_1$ defines a threshold for k_1, $k_1^0 = \dfrac{RW_0}{R-(1-\tau)q}$. If the demand for capital is below this threshold firms will always be able to attain the first-best (FB) case.

4.2.4 Bank's Learning Process

In order to determine the loan's interest rate R_1 the bank has to estimate the probability that the firm may default. This probability will depend on the probability distribution of the prediction θ_1, and therefore, the bank has to learn about θ in a similar way as the firm learns about it.

N is the number of realizations of θ_t observed by the bank and it is assumed to be equal to the number of years the bank has had a relationship with the firm. In other words, I am assuming that the bank can only observe θ_t if it has a relationship with a firm, and therefore, *N* also measures the strength of the relationship between the bank and the firm at time $t = 0$.

It is important to notice that by construction, *N* cannot be higher than *T*. And only if the bank observes θ_t since the moment the firm is born, then $N = T$. If for any reason, the relationship between the bank and the firm starts later than when the firm is born, then $T > N$.

Using information set $\Omega_B = \{\theta_0, ..., \theta_{-N}\}$ the bank can also predict the value of θ_t at $t = 1$. Denoting the future (unobserved) value of θ_t at $t = 1$ as θ_1, we have that $\theta_1 \sim N(\theta_B - \sigma_B^2/2, \sigma_B^2)$ with θ_B given by:

(12)
$$\theta_B = \frac{\hat{\theta}(\sigma^2/N)^{-1} + (\theta_{0B})(\sigma_{0B}^2)^{-1}}{(\sigma^2/N)^{-1} + (\sigma_{0B}^2)^{-1}}$$

and a variance equal to

(13)
$$\sigma_B^2 = \text{var}(\theta_1 \mid \Omega_B) = \left(\frac{1}{(\sigma^2/N)^{-1} + (\sigma_{0B}^2)^{-1}} \right) + \sigma^2$$

Finally, given that θ_1 is normally distributed, we have that $e^{\theta_1} = \Theta_1$ is log-normally distributed with an expected value given by[9]:

(14)
$$\tilde{\Theta}_{1B} = E_0(e^{\theta_1} \mid \Omega_B) = \exp(\theta_B)$$

[9] And a variance given by $\text{var}(e^{\theta_1} \mid \Omega_F) = \left[\exp(\sigma_B^2) - 1\right]\exp(2\theta_B)$

4.2.5 Probability of Default and Loan's Interest Rate R_1

I will denote by $\overline{\Theta}_1$ (or $\overline{\theta}_1$)[10] the default point, such that for any $\Theta_1 < \overline{\Theta}_1$ the firm will not be able to repay the debt, i.e. $\Theta_1 k_1^{\beta} + (1-\tau)qk_1 < R_1 B_1,\ \forall \Theta_1 < \overline{\Theta}_1$. More specifically, $\overline{\Theta}_1$ is the realization of Θ_1 that makes that the firm's value of equity at the end of $t = 1$ is equal to zero. Therefore, the default point $\overline{\Theta}_1$ is implicitly defined by:

(15)
$$e^{\overline{\theta}_1} k_1^{\beta} + (1-\tau)qk_1 - R_1 B_1 = \overline{\Theta}_1 k_1^{\beta} + (1-\tau)qk_1 - R_1 B_1 = 0$$

The default point determines how risky a firm is. Higher default points indicate higher default probabilities, whether smaller default points indicate lower default probabilities. Therefore, $\overline{\Theta}_1$ is a sufficient statistic for characterizing the riskiness of a firm.

According to equation (15), the default point $\overline{\Theta}_1$ can be interpreted as the exact level of profitability that a firm needs to reach in order to be able to repay the amount of debt needed to finance the level of capital k_1, if it has to pay interest rate R_1. Similarly, $\overline{\Theta}_1 k_1^{\beta}$ is the exact amount of revenues that a firm needs to reach to be able to repay its debt at the required interest rate.

For estimating $\overline{\Theta}_1$ the bank also has to determine simultaneously the value of the loan's interest rate R_1, which is implicitly defined according to the following rule:

$$\int_{\overline{\Theta}_1}^{\infty} R_1 B_1 \phi_B d\Theta_1 + \int_0^{\overline{\Theta}_1} \left[(1-\tau)qk_1 + \Theta_1 k_1^{\beta}\right] \phi_B d\Theta_1 = RB_1,$$

that can also be expressed as:

(16)
$$(1-\Phi_B)R_1 B_1 + \Phi_B (1-\tau)qk_1 + (\tilde{\Theta}_{1B} \mid D)k_1^{\beta} = RB_1$$

Where ϕ_B is the probability density function according to the bank's information, Φ_B is the bank's estimated probability of default, and $(\tilde{\Theta}_{1B} \mid D)k_1^{\beta}$ is the partial expectation of future revenues below the default point, according to bank's information[11]. $(\tilde{\Theta}_{1B} \mid D)k_1^{\beta}$ is therefore, the value of the firm's future revenues that the bank expects to receive if the firm defaults. RB_1 is the secure income for the bank when debt is fully collateralized.

Equation (16) simply equals the safe repayment in the riskless case (RB_1) with the expected return in the risky case. The firm repays the debt with probability $(1-\Phi_B)$, and in that case

---

[10] $\overline{\theta}_1 = \ln(\overline{\Theta}_1)$

[11] $\tilde{\Theta}_{1B} = \int_0^{\infty} \Theta_1 \phi_B d\Theta_1 = \int_0^{\overline{\Theta}_1} \Theta_1 \phi_B d\Theta_1 + \int_{\overline{\Theta}_1}^{\infty} \Theta_1 \phi_B d\Theta_1 = (\tilde{\Theta}_{1B} \mid D) + (\tilde{\Theta}_{1B} \mid S)$. In other words, the expected value of Θ_1 could be divided into two partial expectations, the partial expectation below the default point and the partial expectation above the default point.

the bank receives R_1B_1. If the firm defaults, the bank receives the residual value of assets $(1-\tau)qk_1$ with probability Φ_B, plus the expected value of revenues that are left on the firm $(\tilde{\Theta}_{1B} \mid D)k_1^{\beta}$.

Using (16) to replace R_1 into (15), the default point $\overline{\Theta}_1$ is implicitly defined by:

(17) $$\overline{\Theta}_1 k_1^{\beta} - \frac{RB_1 - qk_1}{(1-\Phi_B)} + \frac{(\tilde{\Theta}_{1B} \mid D)k_1^{\beta}}{(1-\Phi_B)} - \frac{\tau q k_1}{(1-\Phi_B)} = 0$$

It is important to notice that the default point $\overline{\Theta}_1$ defined by equation (17) is equal for both the firm and the bank, given that for any given k_1 and after observing R_1, the firm estimates Φ_F based on equation (17), which does not include the firm's probability distribution function for Θ_1, i.e. equation (17) does not include Φ_F. In other words, the default point $\overline{\Theta}_1$ is determined based only on the possible realizations of Θ_1 and on Φ_B (for a given k_1).

Proposition 1*.*

There is a function $\overline{\Theta}_1(k_1) = \overline{\Theta}_1$ *and its partial derivative with respect to k_1 is given by:*

$$\frac{\partial \overline{\Theta}_1}{\partial k_1} > 0 \text{ for } \forall k_1 > k_1^0 \text{ and If } W_0 \geq 0$$

Proof: Appendix A. 3

This proposition says that when capital is financed by risky debt, the higher the amount of capital the more risky is debt. This is however, only guaranteed if $W_0 \geq 0$. If $W_0<0$, the sign of $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ could be sometimes negative and sometimes positive. Most of the propositions are proved by restricting the analysis to the case of positive levels of internal funds. In a separate section of this chapter, I analyze the results in the case that internal funds are negative.

Proposition 2*.*

The default point $\overline{\Theta}_1$ is higher (investment is riskier) the smaller the initial equity (W_0), the smaller the firm's profitability expected by the bank (θ_B or the mean of the distribution), and the higher the bankruptcy costs (τ).

$$\frac{\partial \overline{\Theta}_1}{\partial W_0} < 0, \quad \frac{\partial \overline{\Theta}_1}{\partial \theta_B} < 0, \quad \frac{\partial \overline{\Theta}_1}{\partial \tau} > 0$$

Proof: Appendix A. 3

Corollary 1 *(for propositions 1 and 2).*

The loan's interest rate R_1 and the total loan's repayment R_1B_1 are strictly increasing in the amount of capital k_1 and strictly decreasing in the amount of internal funds (equity) W_0, i.e.

$$\frac{\partial R_1}{\partial k_1} > 0 \; and \; \frac{\partial R_1 B_1}{\partial k_1} > 0, \; \forall k_1 > k_1{}^0$$

$$\frac{\partial R_1}{\partial W_0} < 0 \; and \; \frac{\partial R_1 B_1}{\partial W_0} < 0, \; \forall k_1 > k_1{}^0$$

The total repayment R_1B_1 is a convex function of the amount of capital k_1 (its second derivative with respect to k_1 is positive), i.e.

$$\frac{\partial^2 (R_1 B_1)}{\partial k_1{}^2} > 0, \forall k_1 > k_1{}^0$$

Proof: Appendix A. 3

Proposition 3.
The value of the firm at t = 1under interest rate R_1 is a concave function of the level of investment k_1, i.e. its second derivative with respect to k_1 is always negative.

Proof: Appendix A. 3

Proposition 4.

$$\frac{d\overline{\Theta}_1}{d\sigma_B} > 0 \;,$$

Proof: Appendix A. 3

This proposition shows that the estimated default point is higher the higher the variance of prediction.

4.2.6 Investment Decision with Asymmetric Learning and Bankruptcy Costs

The expected value of firms facing interest rate R_1 is equal to:

(18) $$\mathrm{E}_0(V_0) = \frac{1}{R} E_0\left[\left(\int_{\overline{\Theta}_1}^{\infty} (\Theta_1 k_1{}^{\beta} + qk_1 - R_1 B_1)\phi_F \, d\Theta_1 + \int_0^{\overline{\Theta}_1} (0)\phi_F \, d\Theta_1\right) \mid \Omega_F\right] =$$

$$= \frac{1}{R}\left[(\tilde{\Theta}_{1F} \mid S)k_1^{\beta} + qk_1(1-\Phi_F) - R_1 B_1 (1-\Phi_F)\right]$$

where ϕ_F is the firm's probability density function for Θ_1, Φ_F is the firm's estimated probability of default (the cumulative distribution). $(\tilde{\Theta}_{1F} \mid S)k_1^{\beta}$ is the partial expectation of future revenues in the case the firm "succeeds" in repaying the debt (the expected value above the default point) based on the firm's information.

Using again (16) into (18) to eliminate R_1 we have that the expected value of equity at $t = 0$ is equal to:

(19) $$E_0(V_0) = \frac{1}{R}\left[\left((\tilde{\Theta}_{1F} \mid S) + \frac{(1-\Phi_F)}{(1-\Phi_B)}(\tilde{\Theta}_{1B} \mid D)\right)k_1^{\beta} - \frac{(1-\Phi_F)}{(1-\Phi_B)}\left((RB_1 - qk_1) + \tau\Phi_B qk_1\right)\right]$$

In Appendix A. 4 (after some algebra) I show that the first order condition with respect to k_1 is given by:

(20) $$\left((\tilde{\Theta}_{1F} \mid S) + \frac{(1-\Phi_F)}{(1-\Phi_B)}(\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{\beta-1} = \left((R-q) + \tau q\Phi_B\right)\frac{(1-\Phi_F)}{(1-\Phi_B)} + \phi_F \frac{\partial \overline{\Theta}_1}{\partial k_1}\tau q k_1$$

(In the Appendix A. 4 it is shown that this is exactly the same as deriving equation (18) with respect to k_1).

If the bank has less information than the firm, its perception of risk will be higher, because the bank's uncertainty about the true firm quality is higher and therefore, the bank's prediction of future profits is less precise.

Therefore, we can summarize the difference in the bank's and firm's information in the parameters T and N. If $T = N$, there is symmetry of information and $\sigma_F^2 = \sigma_B^2$. If $T > N$ there is asymmetry of information and the variance of prediction of the bank is higher than the firm's variance $\sigma_F^2 > \sigma_B^2$ [12].

Variances of firm and bank are both decreasing functions of T and N (the age and the duration of relationship). Therefore, we can express the "average" distributions of firm and bank as $\theta_1 \sim N(\theta - \sigma_F^2/2, \sigma_F^2)$ and $\theta_1 \sim N(\theta - \sigma_B^2/2, \sigma_B^2)$ respectively, meanwhile Θ_1 is log-normally distributed according to the same parameters.

It is important to understand the different effects that the evolution of the variance has on the cumulative distribution function and the including the partial expectations of Θ_1 .

First, if $\sigma_B > \sigma_F$, the estimated probabilities of default of bank and firm would be different given that the cumulative distribution function would be different, i.e. $\Phi_B \neq \Phi_F$. Whether this difference is positive or negative depends on the value of the default point $\overline{\theta}_1$. Second, the partial expectations below the default point would also be different, i.e. $(\tilde{\Theta}_{1B} \mid D) \neq (\tilde{\Theta}_{1F} \mid D)$. The sign of this difference also depends on the value of the default point $\overline{\theta}_1$.

[12] There could be also be asymmetry in the *prior* information, which would be reflected in the *prior* variance.

Second, as both firm and bank accumulate more information (as they learn) both variances decreases (whatever the initial difference). As both variance decreases the cumulative distribution functions change, as well as the partial expectations. Therefore, even in the case that both firm and bank have the same information (symmetric case) the estimated probability of default, and the expected value of the firm would change as the firm and the bank learn, as both $T \to \infty$ and $N \to \infty$.

In general, how do these functions change as the variance change may depend on the value of the default point $\overline{\Theta}_1$, and thus, they may depend on the value of W_0 and B_1.

4.3 Relationship between Investment and Internal Funds

For deriving the sign of the relationship between investment and internal funds, in the case of asymmetry of information and positive bankruptcy costs, we need to use the first order condition for the optimal level of capital and derive implicitly with respect to internal funds. From now on I denote the optimal level of capital defined by the first order condition as k_1^*.

It is important to remind that in the first-best case (symmetry and $\tau = 0$) we have that: $\frac{\partial k_1^*}{\partial W_0} = 0$, which means that in the optimal case, the optimal level of capital does not change with the level of internal funds (wealth).

I will use here the fact that the expected value of the firm is a concave function and therefore that $\frac{\partial V(k_1^*, \overline{\Theta}_1)}{\partial^2 k_1^*} < 0$.

4.3.1 First Case: Asymmetry without Bankruptcy Costs

In the case of asymmetry of information without bankruptcy costs, the first order condition for optimal investment is given by:

(21)
$$\left((\tilde{\Theta}_{1F} \mid S) + \frac{(1-\Phi_F)}{(1-\Phi_B)} (\tilde{\Theta}_{1B} \mid D) \right) \beta k_1^{\beta-1} = (R-q) \frac{(1-\Phi_F)}{(1-\Phi_B)}$$

Multiplying both sides by $\frac{(1-\Phi_B)}{(1-\Phi_F)}$ this can also be written as:

$$\textbf{(22)} \qquad \left((\tilde{\Theta}_{1F} \mid S) \frac{(1-\Phi_B)}{(1-\Phi_F)} + (\tilde{\Theta}_{1B} \mid D) \right) \beta k_1^{\beta-1} = (R-q)$$

In the symmetric case we would have that $(\tilde{\Theta}_{1B} \mid D) = (\tilde{\Theta}_{1F} \mid D)$ and that $\Phi_B = \Phi_F$ and therefore (22) would be equivalent to the first-best FOC,

$$\tilde{\Theta}_{1F} \beta k_1^{\beta-1} = \tilde{\Theta}_{1B} \beta k_1^{\beta-1} = (R-q)$$

To simplify the analysis, I will use the definition of the expected value given default for a log-normal distribution, according to which we have that $(\tilde{\Theta}_{1B} \mid D) = \tilde{\Theta}_{1B} \Phi_B \left((\exp(\overline{\theta}_1 - \sigma_B) \right)$. This is saying that the expected value given default can be expressed as the expected value $\tilde{\Theta}_{1B}$ multiplied by the probability of the default point minus the standard deviation. According to this definition it is clear that $\Phi_B[(\exp(\overline{\theta}_1 - \sigma_B)] < \Phi_B(\exp(\overline{\theta}_1)) = \Phi_B$.

Similarly, according to the definition of the expected value given success, we have that $(\tilde{\Theta}_{1B} \mid S) = \tilde{\Theta}_{1B} \left[1 - \Phi_B \left((\exp(\overline{\theta}_1 - \sigma_B) \right) \right]$ and that $\left(1 - \Phi_B[(\exp(\overline{\theta}_1 - \sigma_B)] \right) > (1-\Phi_B)$.

These definitions are the same for both bank and firm, i.e. we can change the subscript *B* by *F* to write down the same functions for the firm. I also denote $\Phi_B \left((\exp(\overline{\theta}_1 - \sigma_B) \right)$ as Φ_B^{EVD}. Therefore, I can also denote $(\tilde{\Theta}_{1B} \mid D)$ as $\tilde{\Theta}_{1B} \Phi_B^{EVD}$, and $(\tilde{\Theta}_{1B} \mid S)$ as $\tilde{\Theta}_{1B} \left[1 - \Phi_B^{EVD} \right]$, changing the subscript *B* by *F* to denote the firm.

Replacing this notation into equation (22) and using that $\tilde{\Theta}_{1B} = \tilde{\Theta}_{1F}$ we get the following:

$$\textbf{(23)} \qquad \left((1-\Phi_F^{EVD}) \frac{(1-\Phi_B)}{(1-\Phi_F)} + \Phi_B^{EVD} \right) \beta \tilde{\Theta}_{1F} k_1^{\beta-1} = (R-q)$$

Here is important to remind that equation (22) and equation (23) are exactly the same, only with a different notation. Rearranging terms in equation (23) we also can write it as:

$$\textbf{(24)} \qquad \beta \tilde{\Theta}_{1F} k_1^{\beta-1} = (R-q) \frac{1}{\left((1-\Phi_F^{EVD}) \frac{(1-\Phi_B)}{(1-\Phi_F)} + \Phi_B^{EVD} \right)}$$

Equations (21) to (24) can help us to understand the intuition of the u-shaped investment relation between investment and wealth (or leverage, or risk). Comparing these two last expressions with the FOC in the "first-best" case we can see that whenever the expression $\left((1-\Phi_F^{EVD}) \frac{(1-\Phi_B)}{(1-\Phi_F)} + \Phi_B^{EVD} \right)$ is lower than one, investment would be smaller in the asymmetric case because either the marginal product of capital would be lower, or the marginal cost would be higher.

In order to show that investment is always lower in the asymmetric case we need to prove that $\left((1-\Phi_F^{EVD})\dfrac{(1-\Phi_B)}{(1-\Phi_F)} + \Phi_B^{EVD} \right)$ is always smaller than 1, or equivalently, that

$$\left((1-\Phi_F^{EVD})\frac{(1-\Phi_B)}{(1-\Phi_F)} + \Phi_B^{EVD} \right) < \left((1-\Phi_F^{EVD})\frac{(1-\Phi_F)}{(1-\Phi_F)} + \Phi_F^{EVD} \right)$$

Thus, investment is lower in the asymmetric and the symmetric cases when the difference between the two sides of the last equation is smaller than zero, and they are equal when the difference is equal to zero, i.e. whenever:

$$\left((1-\Phi_F^{EVD})\frac{(1-\Phi_B)}{(1-\Phi_F)} + \Phi_B^{EVD} \right) - \left((1-\Phi_F^{EVD})\frac{(1-\Phi_F)}{(1-\Phi_F)} + \Phi_F^{EVD} \right) = 0$$

Rearranging terms we get that:

$$\left((1-\Phi_F^{EVD})\left[\frac{(1-\Phi_B)}{(1-\Phi_F)} - \frac{(1-\Phi_F)}{(1-\Phi_F)} \right] + (\Phi_B^{EVD} - \Phi_F^{EVD}) \right) = 0$$

(25) $$\left(\frac{(1-\Phi_F^{EVD})}{(1-\Phi_F)} \left[(1-\Phi_B) - (1-\Phi_F) \right] + (\Phi_B^{EVD} - \Phi_F^{EVD}) \right) = 0$$

The expression shows that investment could be smaller in the asymmetric case because the bank estimates a smaller probability of repayment, i.e. because $(1-\Phi_B) < (1-\Phi_F)$; or because the bank expects smaller firm's revenue if the firms defaults, i.e. $(\tilde{\Theta}_{1B} \mid D) < (\tilde{\Theta}_{1F} \mid D)$, which is exactly equivalent to $\Phi_B^{EVD} < \Phi_F^{EVD}$.

In Figure 8 we can see how the expected probabilities of repayment $(1-\Phi_B)$ and $(1-\Phi_F)$ differ when the bank has a higher variance (higher uncertainty) because $T > N$, as a function of the default point. In the Figure we can see that for the lower levels of a default point (highest W_0) the probability of repayment estimated by the bank is lower (left part of the graph). The opposite happens when the default point is high (lowest W_0).

We can also see in Figure 8 that something similar happens with respect to the expected value given default $(\tilde{\Theta}_{1B} \mid D)$ and $(\tilde{\Theta}_{1F} \mid D)$. In this case, the bank's expected value is higher for the lowest values of risk (highest W_0) and lower for the highest values of risk (lowest W_0, left side of the graph).

In summary, with asymmetric information there are two counteracting effects that depend on the level of risk and thus on W_0. When risk is low, the marginal cost (product) should be higher (lower) because $(1-\Phi_B) < (1-\Phi_F)$, but at the same, there is a counteracting effect because $(\tilde{\Theta}_{1B} \mid D) > (\tilde{\Theta}_{1F} \mid D)$. On the other hand, when risk is high investment should be higher because $(1-\Phi_B) > (1-\Phi_F)$, but at the same time, there is a counteracting effect because $(\tilde{\Theta}_{1B} \mid D) < (\tilde{\Theta}_{1F} \mid D)$.

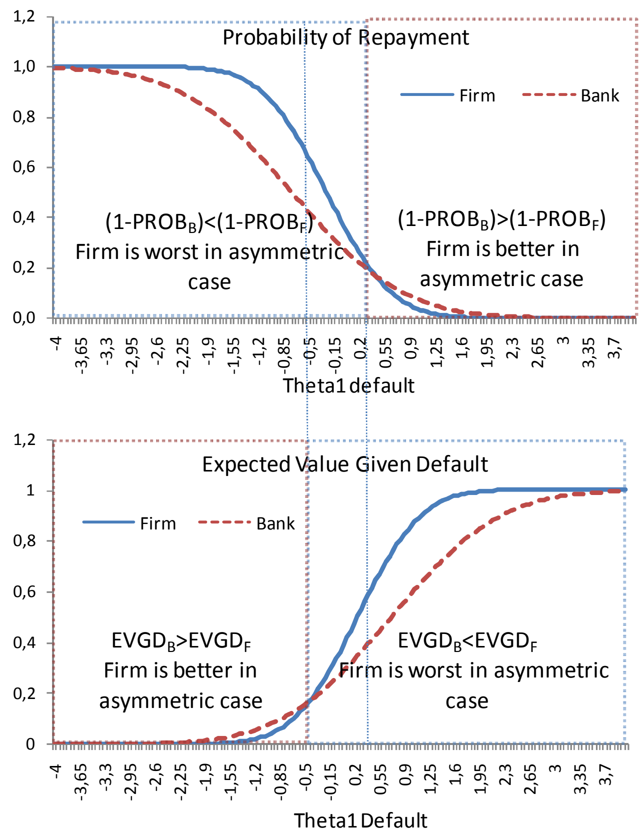


Figure 8 Probability of Repayment and Expected Value given Default according to the level of default risk.

Albeit these counteracting effects of asymmetry, we can see graphically in Figure 9 that the effect of a higher variance always translates into a higher (lower) marginal cost (product) of capital. Furthermore, this effect is larger for the medium levels of risk, which translates into a u-shaped relationship between capital and internal funds.

Figure 9 show three functions. The first one shows the difference between the estimated probabilities of repayment of bank and firm $[(1-\Phi_B)-(1-\Phi_F)]$ multiplied by the term $\frac{\Phi_F^{EVD}}{(1-\Phi_F)}$, which corresponds to the first term in equation (25). This function is obviously negative when the bank assigns a lower probability of repayment, i.e. when $(1-\Phi_B)<(1-\Phi_F)$, and positive the other way around.

The second function is the difference between the expected value given default estimated by the bank $(\tilde{\Theta}_{1B} \mid D)$ and the expected value given default estimated by the firm $(\tilde{\Theta}_{1F} \mid D)$, which is equivalent to the difference between Φ_B^{EVD} and Φ_F^{EVD} (the second term in equation (25)).

The third function called "Difference in Marg. Cost/Prod." is the graph of equation (25). When the graph is negative the marginal cost (product) in the asymmetric case is larger (smaller) than in the symmetric one, and thus investment in the asymmetric case should be lower. It is clear in Figure 9 that the function Difference is always negative and that the smaller values are around the medium levels of risk. This is translated into a u-shaped function of investment with respect to internal funds (risk).

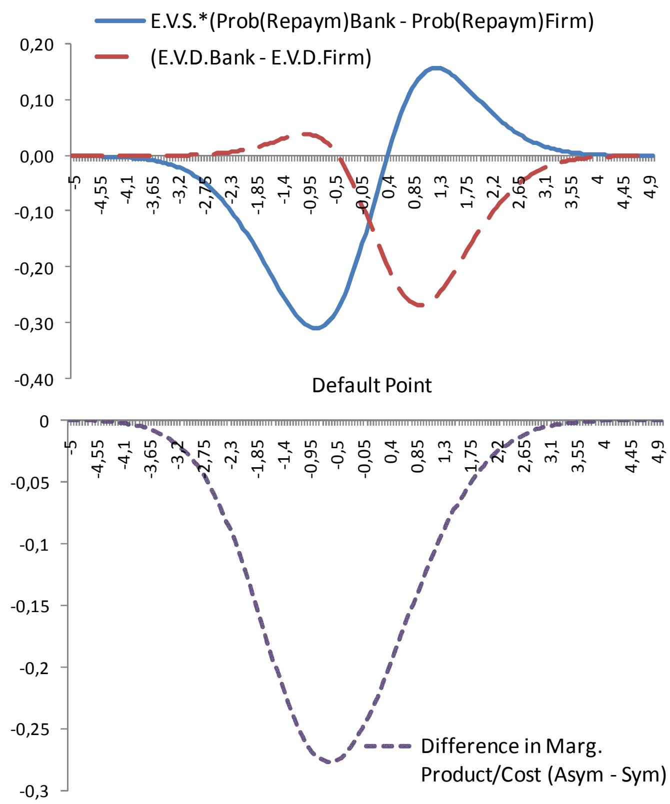


Figure 9 Difference in functions that composed marginal cost (product). A negative Difference implies a larger (smaller) marginal cost (product) in the asymmetric case.

In *Proposition 5* I formally show that the effect of a bank's higher variance always translates into a higher (lower) marginal cost (product) of investment and therefore into a lower level of capital k_1.

Proposition 5.

The sum of the effects of a higher variance in the probability of repayment and the expected value given default is always "negative", meaning that the marginal product in equation (22) is always smaller in the asymmetric case than in the symmetric one, or that equation (25) is always negative.

Proof: Appendix A. 3

To formally prove that the relationship between investment and internal funds (risk) is non-monotonic as we see in Figure 9, we also need to derive the FOC given by equation (22) with respect to W_0, in order to derive the sign of $\frac{\partial k_1^*}{\partial W_0}$.

Furthermore, to derive an expression that can be tested with the data, I take logs at both sides of equation (22) and, isolating the log of k_1 we get the following expression for investment capital:

(26)
$$\ln k_1^* = -\frac{\ln\beta}{(\beta-1)} - \frac{\ln\tilde{\Theta}_{1F}}{(\beta-1)} + \frac{\ln(R-q)}{(\beta-1)} - \frac{\ln\left((1-\Phi_F^{EVD})\frac{(1-\Phi_B)}{(1-\Phi_F)} + \Phi_B^{EVD}\right)}{(\beta-1)}$$

Proposition 6

The derivative of the optimal level of investment with respect to internal funds is given by:

$$\frac{\partial k_1^*}{\partial W_0} = \frac{\frac{\partial\overline{\Theta}_1}{\partial W_0}\left(\overline{\Theta}_1 - \frac{(\tilde{\Theta}_{1F} \mid S)}{(1-\Phi_F)}\right)\left(\phi_B - \phi_F\frac{(1-\Phi_B)}{(1-\Phi_F)}\right)\beta k_1^{\beta-1}}{-\frac{\partial V(k_1^*,\overline{\Theta}_1)}{\partial^2 k_1^*}}$$

The sign of $\frac{\partial k_1^*}{\partial W_0}$ *is positive whenever* $\left(\phi_B - \phi_F\frac{(1-\Phi_B)}{(1-\Phi_F)}\right) > 0$.

The sign of $\frac{\partial k_1^*}{\partial W_0}$ *is negative whenever* $\left(\phi_B - \phi_F\frac{(1-\Phi_B)}{(1-\Phi_F)}\right) < 0$.

Proof: *Appendix A. 3*

Corollary for Proposition 6

The $\ln(k_1)$ depends on the function $\left((1-\Phi_F^{EVD})\frac{(1-\Phi_B)}{(1-\Phi_F)} + \Phi_B^{EVD}\right)$ *which is a non-monotonic function of the level of internal funds W_0.*

Proof: *The proof of this corollary is straightforward since the derivative of* $\left((1-\Phi_F^{EVD})\frac{(1-\Phi_B)}{(1-\Phi_F)} + \Phi_B^{EVD}\right)$ *is exactly the same as the numerator of* $\frac{\partial k_1^*}{\partial W_0}$ *without the term* $\beta k_1^{\beta-1}$.

In Appendix A. 3 it is shown that the term $\left(\phi_B - \phi_F\frac{(1-\Phi_B)}{(1-\Phi_F)}\right)$ is more likely to be negative when the default point is high (low W_0) and is more likely to be positive when the default point is low (high W_0).

This again implies that when W_0 is low and therefore $\overline{\Theta}_1$ is high, the relationship between investment and wealth should be negative, and when W_0 is high and therefore $\overline{\Theta}_1$ is low, the relationship between investment and wealth should be positive.

The rationale for the u-shaped relationship in the case that the bank has a shorter history of information than the firm is the following: Any disagreement between a bank and a firm matters most when repayment or default are equally likely, and it matters less when any of them are almost certain.

Intuitively, if the firm itself is not sure if it will default or if it will be able to repay its debt because it predicts that both situations are equally likely, any difference in the creditor's information could make a large difference in its prediction about the likelihood of those events, compared to the firm's beliefs. On the other hand, if either default or repayment are almost certain, the predictions from both sides are going to be quite similar independently of the set of information that they are using.

Therefore, the effect of a disagreement between a creditor and a firm caused by the creditor's higher uncertainty (variance) depends on the level of internal funds: it is more important for intermediate levels of risk (intermediate levels of internal funds) and it is less important for high and low levels of risk (low and high levels of internal funds).

In the model, this translates into the following: When bank and firm share the same information they predict the same probability of default, and therefore, they both agree on the fair cost of debt independently of how risky is investment, i.e. independently of the level of internal funds. However, with asymmetry of information the bank may estimate a lower probability of repayment, or the bank may expect to recover a lower value of revenues in the case the firm defaults (compared to what the firm expects to transfer to the bank in that case). When any of those situations occurs (or both of them) the bank charges the firm a higher interest rate than what the firm considers fair. Therefore, the marginal cost of investment is higher than in the symmetric case and the firm must decrease its level of capital to increase its marginal revenue, to compensate for the higher marginal cost. In Figure 10 we can see an example of the relationship between capital (y-axis) and internal funds (x-axis) as described by equation (23), and the log of capital versus the log of $\left((1-\Phi_F^{EVD})\frac{(1-\Phi_B)}{(1-\Phi_F)}+\Phi_B^{EVD}\right)$ as described by equation (26). The example assumes that T=3 and N=0.

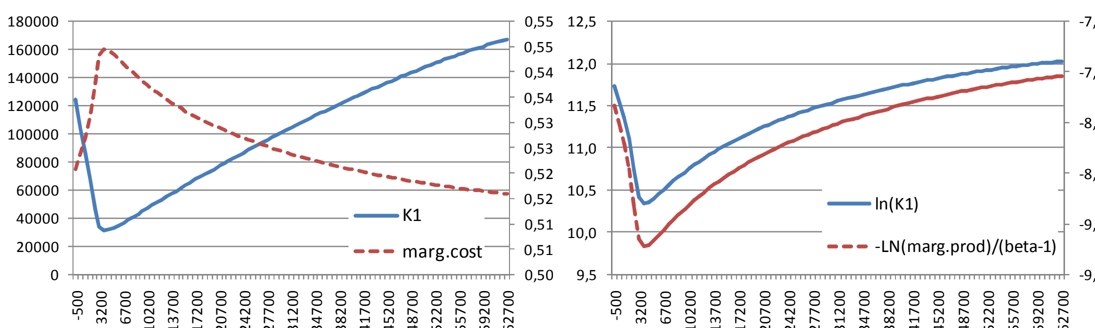


Figure 10. Optimal capital vs. Marginal product of capital

4.3.2 Second Case: Symmetry of Information but Positive Bankruptcy Costs

The optimal level of capital in this case is given by the first order condition:

(27)
$$\beta\tilde{\Theta}_{1F}k_1^{*\beta-1} = (R-q) + \tau q\Phi_B + \tau q k_1 \phi_B \frac{\partial\overline{\Theta}_1}{\partial k_1}$$

Or

(28)
$$\frac{\partial V(k_1^*, \overline{\Theta}_1)}{\partial k_1} = \beta\tilde{\Theta}_{1F}k_1^{*\beta-1} - (R-q) - \tau q\Phi_B - \tau q k_1 \phi_B \frac{\partial\overline{\Theta}_1}{\partial k_1} = 0$$

Taking logs at both sides of equation (28) and isolating the log of k_1 we get the following (using the properties of the logarithm of a sum):

(29)
$$\ln k_1^* = -\frac{\ln\beta}{(\beta-1)} + \frac{\ln(R-q)}{(\beta-1)} - \frac{\ln\tilde{\Theta}_{1F}}{(\beta-1)} + (1/(\beta-1))\ln\left(1 + \frac{\tau q\left(\Phi_B + \phi_B k_1 \frac{\partial\overline{\Theta}_1}{\partial k_1}\right)}{(R-q)}\right)$$

In this section, I restrict the analysis to the case of $W_0 \geq 0$ and $\frac{\partial\overline{\Theta}_1}{\partial k_1} > 0$. Later on in the chapter I analyze the case of $W_0 \geq 0$ and the possibility of $\frac{\partial\overline{\Theta}_1}{\partial k_1}$ being negative and I show that it does not affect the results.

The next proposition shows that the optimal level of investment in the case of symmetry but positive bankruptcy costs could be a non-monotonic, u-shaped function of internal funds.

Proposition 7

The derivative of the optimal level of investment with respect to internal funds is given by:

(30)
$$\frac{\partial k_1^*}{\partial W_0} = \frac{-\frac{\partial\overline{\Theta}_1}{\partial W_0}\tau q\phi_B\left[(1-\beta) + \frac{\partial\overline{\Theta}_1}{\partial k_1^*}k_1^*\left[\frac{(-\overline{\theta}_1 - 3\sigma_B^2/2 + \theta_B)(1-\Phi_B) + \phi_B\overline{\Theta}_1\sigma_B^2}{\overline{\Theta}_1\sigma_B^2(1-\Phi_B)}\right]\right]}{-\frac{\partial V(k_1^*, \overline{\Theta}_1)}{\partial^2 k_1^*}}$$

The sign of $\frac{\partial k_1^*}{\partial W_0}$ *is positive whenever* $\overline{\theta}_1 < \theta_B - 3\sigma_B^2/2$. *Or in other words, the relationship between investment and wealth is always positive for values of the default point smaller than the mode of the distribution* $\theta_B - 3\sigma_B^2/2$ *(less risky points). This means that the higher internal funds are, the lower the default point and the more likely that the relation is positive.*

The sign of $\frac{\partial k_1^*}{\partial W_0}$ *could be negative if* $\bar{\theta}_1 > \theta_B - 3\sigma_B^2/2$ *, i.e. for those points higher than the mode of the distribution (more risky points). This means that the lower internal funds are, the higher the default point and the more likely that the relation is negative.*

The sign is more likely to be negative the higher the variance and the more skewed the distribution is.

Proof: Appendix A. 3

Corollary for Proposition 7

The ln(k₁) depends on the function $\left(\tau q\Phi_B + \tau q k_1 \phi_B \frac{\partial \overline{\Theta}_1}{\partial k_1}\right)$ *which is a non-monotonic function of the level of internal funds.*

Proof: The proof of the corollary is straightforward since the derivative of the function $\left(\tau q\Phi_B + \tau q k_1 \phi_B \frac{\partial \overline{\Theta}_1}{\partial k_1}\right)$ is exactly the same as the numerator of $\frac{\partial k_1^*}{\partial W_0}$.

We may get the intuition of the u-shaped relationship if we look at the marginal cost of capital given by the first order condition which is equal to $(R-q) + \tau q\Phi_B + \tau q k_1 \phi_B \frac{\partial \overline{\Theta}_1}{\partial k_1}$.

The marginal cost of a unit of capital is given by three terms. The first one is the marginal cost of a unit of risk-free capital (*R-q*). The second term accounts for the fact that for each marginal unit of capital k_1, a percentage $\tau q\Phi_B$ is lost. The third term accounts for the fact that for each marginal unit of capital (keeping constant W_0) the probability of default increases and thus the loss per unit is higher in expected terms. The latter effect depends on the marginal increase in the probability of default due to a marginal increase in k_1 (given by $\frac{\partial \Phi_B}{\partial k_1} = \phi_B \frac{\partial \overline{\Theta}_1}{\partial k_1}$).

We have to bear in mind that in the model the probability of default changes when leverage increases, and since $B_1 = k_1 - W_0$, an increase in the demand for debt could be due to a change in the level of capital (keeping constant internal funds), due to a change in internal funds (keeping constant the level of capital), or from a change in both of them.

In both cases we expect that an increase in leverage translates into an increase in default risk and default probability. In other words we expect that both $\frac{\partial \overline{\Theta}_1}{\partial k_1} > 0$ and that $\frac{\partial \overline{\Theta}_1}{\partial W_0} < 0$ and therefore we expect that $\frac{\partial \Phi_B}{\partial k_1} > 0$ and $\frac{\partial \Phi_B}{\partial W_0} < 0$.

But there is no reason to expect that the cross derivative of Φ_B with respect to W_0 and k_1 should be positive, or that the second derivative of Φ_B with respect to W_0 should be positive.

In fact what I show is that both of those higher derivatives have a different sign depending on the level of risk (and thus depending on W_0 itself).

The reason for this is that any change in the probability of default is determined by its density function ϕ_B which depends on the level of risk in a non monotonic way. The density function of a log-normal distribution is increasing at the lower values, it has a maximum at its mode, which is given by $(\theta_B - 3\sigma_B^2/2)$, and it is decreasing afterwards. This can be seen if we derive the density function with respect to the default point:

$$\frac{\partial \phi_B}{\partial \overline{\Theta}_1} = \phi_B \times \left(\frac{-\overline{\theta}_1 + \theta_B - 3\sigma_B^2/2}{\overline{\Theta}_1 \sigma_B^2} \right)$$

The sign of this derivative is positive for values of the default point lower than the mode of the distribution (lower risk), and is negative for values above the mode (higher risk). This result is common to almost every distribution function, with the exception of the uniform distribution.

This implies that when $\overline{\Theta}_1$ is very low and below the mode of the distribution, an increase in risk (a decrease in W_0) generates only a small increase in the probability of default. However, the change in the probability of default becomes larger the larger is the default point (the lower is W_0) and it reaches a maximum when the default point is equal to the mode of the distribution. The mode is the most frequent value of a distribution.

Thus, when $\overline{\Theta}_1$ is around the mode of the distribution, an increase in risk (a decrease in W_0) generates the highest increase in the probability of default. If the increase in risk (decrease in W_0) is large enough that the default point becomes higher than the mode of the distribution, the effect of a decrease in W_0 starts becoming less and less important. When $\overline{\Theta}_1$ is really high and above the mode of the distribution, an increase in risk (a decrease in W_0) generates again only a small (positive) increase in the probability of default.

It is very important to notice that this result implies that the probability of default and the interest rate always increases with a decrease in internal funds, but the increase is smaller when risk is very high or very low.

A graphical example of this description could be seen in Figure 11.

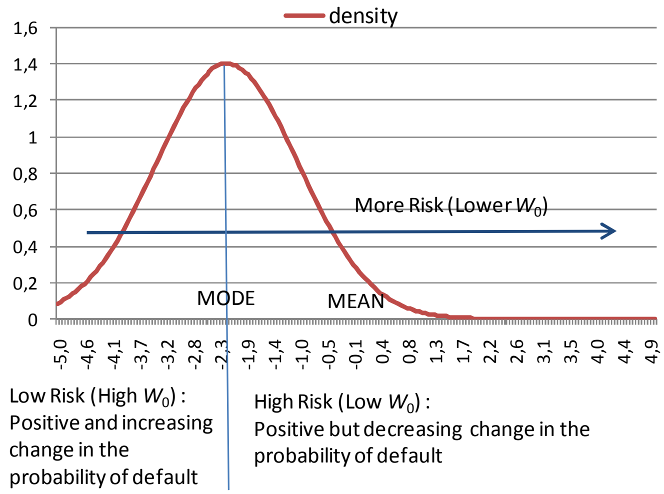


Figure 11

Hence, the rationale for the u-shaped relationship in the case of positive bankruptcy costs is the following: The effect of a marginal decrease in internal funds on the predicted probability of default is higher when repayment or default are similarly likely, and lower when any of them are almost certain.

The intuition in this case is similar to the asymmetric versus symmetric case; if a creditor is uncertain about whether a firm will default or will be able to repay its debt because the creditor believes that both situations are similarly likely, any marginal change in internal funds and in the firm's leverage could largely change the creditor's belief about which event is more likely than the other. On the contrary, if the creditor believes that either default or repayment are almost certain, a marginal change in internal funds cannot change too much its prediction about the relative likelihood of those events.

Therefore, the effect of a marginal change in internal funds and in firm's leverage depends on the level of internal funds: it is more important for intermediate levels of risk (intermediate levels of internal funds) and it is less important for high and low levels of risk (low and high levels of internal funds).

In the model, this translates into the following: when internal funds are higher than the optimal level of investment, the firm does not require external funds and the probability of default is zero. If internal funds decrease a little and the firm has to demand some external funds, the default probability increases from zero to a positive level. As internal funds are high and the default point is very low, a small decrease in internal funds generates only a small increase in the probability of default and in the marginal cost. As the marginal cost of capital increases, the firm has to reduce the level of capital to increase its marginal revenue and to compensate for the increase in the probability of default and the interest rate.

However, the increase in the probability of default becomes larger as internal funds decrease further and the default point increases. Therefore, the decrease in investment has to be larger with every change in internal funds to compensate for the incremental increase in the marginal cost. Such incremental increase in the probability of default reaches a maximum when the default point reaches the mode of the distribution (the most frequent value).

If internal funds decrease enough to take the default point above the mode of the distribution, any further decrease in internal funds generates a smaller increase in the probability of default than before, and therefore, the decrease in investment becomes smaller with every change in internal funds. At this point, the level of capital is very low, and the marginal return per unit of capital is, consequently, very high. When internal funds decrease further enough and the increase in the probability of default per unit of capital is really small, the high marginal return per unit of capital is large enough to compensate for the increase in the marginal cost per unit of capital, and therefore, the firm prefers to increase investment.

The consequence of this is that the total marginal cost does not change monotonically with W_0, and this is the rationale for a non-monotonic relationship between investment and wealth. In Figure 12 and Figure 13 we can see this for an example of a firm with $\tau=0.1$. In Figure 13 we see the relationship between optimal capital and internal funds versus the marginal cost according to equation (27) and also the relationship between the optimal level of capital in logs and the marginal cost in logarithms according to equation (29).

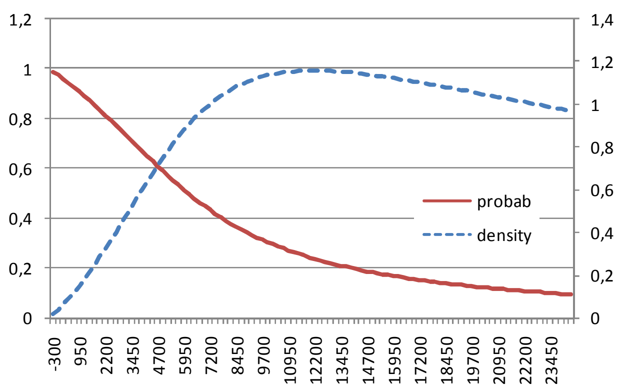


Figure 12. Probability of default (cumulative) and density functions.

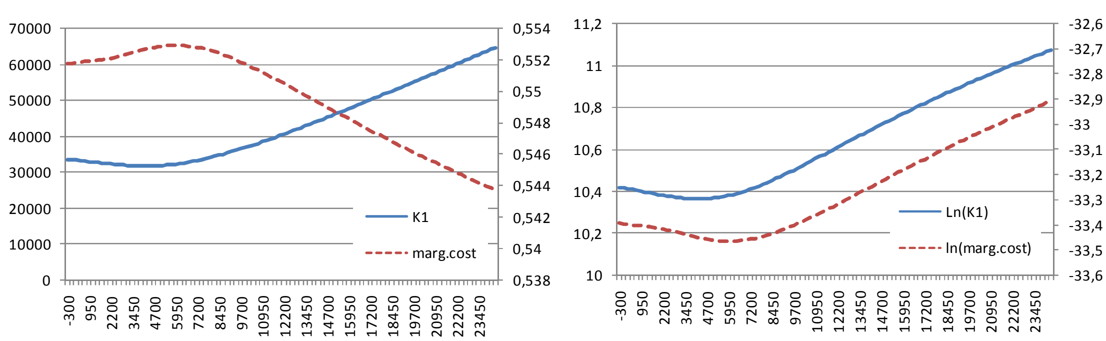


Figure 13 Optimal investment vs. marginal cost

According to these results, it is clear that there may not be a monotonic relationship between investment and internal funds even in the absence of asymmetry of information. In this case, financial constraints in terms of bankruptcy costs may generate a non-monotonic relationship.

In Appendix A. 5 I show the case of Asymmetry together with positive bankruptcy costs ($\tau > 0$).

Another interesting result of this section is that, since the mode of a log-normal distribution depends negatively on the variance of the bank, the higher the variance (uncertainty), the lower the point at which the increase in the probability of default reaches its maximum, and the more likely that a decrease in W_0 may generate only a low increase in the default point. Therefore, even without asymmetry of information, young firms are more likely to show a u-shape investment curve behavior.

4.4 Simulation of Investment Financial-Wealth Sample

As a way to illustrate and evaluate the theoretical predictions of the model about the relationship between investment and internal funds, I perform a simulation exercise using the FOC for capital calibrated according to realistic parameters and values obtained from different samples of firms in order to have an idea of how the u-shaped investment curve should look in an empirical exercise.

Since the most important exogenous variables in the model are internal funds, the age of the firm, the length of the relationship and the relationship between the age and the length (symmetry vs. asymmetry), I replicate the statistical distribution characteristics of these three variables from two different samples of firms in a generated random sample of 3000 observations composed of 600 simulated firms, with 5 year-observations per firm.

I obtain the distribution parameters for the variables Financial Wealth/Assets and Age from the SABI database, and the distribution parameters for the variable Length/Age from the NSSBF 2003. I use the distribution of the variable Length/Age from the NSSBF 2003 database because the SABI database does not have information on the length of bank-firm relationship.

I calibrate the parameters of the model so that the first-best level of capital of the theoretical model is equal to the mean value of the variable Assets in the SABI sample, which is equal to € 904.000. The parameters used are the following: the risk-free real interest rate R is set equal to 1.02, which is the average real return on a one-year U.S. T-bill between 1986 and 2005. The returns-to-scale parameter is set equal to 0.97. The depreciation rate is set to 10% and the percentage of tangible assets is set in 60%, therefore, the parameter q is equal to 50%. The value of τ is set at 10%.

The risk-free interest rate, the returns-to-scale parameter and the depreciation rate come from Caggese (2007). The percentage of tangible assets comes from Shaikh (2004) who suggests that 50-90% of a company's value is derived from its intangible assets and from Blair and Wallman (2001) who estimates that the percentage of the market value of a corporation due to tangible assets has varied from 80% to 20% from 1974 to 1998 (50% on average). The percentage of assets lost after liquidation comes from Andrade and Kaplan (1998) that estimate that the costs of financial distress are 10% of firms' value. Different estimates vary around 3% and 20% of the firm's predistress market value or market value of assets as described by Brealey and Myers (2003).

Since the first-best level is given by $k_1^* = \left[\frac{R-q}{\beta \tilde{\Theta}_{1F}} \right]^{1/(\beta-1)} = 904.000$, the value of $\tilde{\Theta}_{1F}$ must be equal to 0.8089, which implies that θ, the mean of the normal distribution, must be equal to -0.21209. I then generate 3000 random values of the profitability factor θ_1, with a mean equal to -0.21209 and standard deviation of 0.01538, calculated so that a firm whose assets are one standard deviation above the mean in the simulated sample has the same size that a firm that is one standard deviation above the mean in the SABI sample. A firm whose value of assets is one standard deviation above the mean in the SABI sample would have a size of €1.509.713.

I then use the distribution of the ratio of (Financial Wealth/Assets) to calculate 10 different ranges of internal funds by multiplying the defined first-best value of assets by the ratio. The variable financial wealth is defined as the sum of the variables cash in hand, accounts receivables and inventory minus the variable accounts payable. I use the ten percentiles values of the ratio of Financial Wealth to Assets in the SABI sample to calculate ten different ranges of internal funds, which are shown in Table 28.

I generate 300 random values of internal funds for each range. Therefore I generate 300 random values from -353.820 to -126.793, 300 random values from -126.739 to -11.696, etc. according to a uniform distribution inside each range. Therefore I finally have a random sample of 3000 values of internal funds that follows the same statistical distribution of the variable Financial Wealth/Assets in the SABI database. It is important to notice that for the first, lowest range of internal funds, the model is not able to find any solution with a positive value of profits, and similarly for the second lowest range. Therefore, the final sample is smaller than the 3000 total random values of internal funds.

This could be due to the fact that the theoretical model is a one period model that cannot capture the whole dynamics involved in the real data. For instance, the model cannot capture that a firm can repay its debts in several periods of time and not only in one period, as it is implied by the model.

Table 17 Percentiles of Internal Funds according to SABI 2009

WEALTH/ASSETS	MEDIAN ASSETS K1	F.WEALTH W0
-0.391394	904000	-353820
-0.1402579	904000	-126793
-0.0129386	904000	-11696
0.0739801	904000	66878
0.1624594	904000	146863
0.2623722	904000	237184
0.3871375	904000	349972
0.559989	904000	506230
0.8450163	904000	763895
1.103613	904000	997666

I also compute the 4 quartiles of the variable AGE in the NSSBF and generate 400 random values for each of the 5 ranges composed by the 4 quartiles and the minimum and maximum values, as shown in Table 18.

Afterwards, for each range of the variable Age I compute the quartiles of the variable Length/Age according to the sample of NSSBF 2003. Therefore, I compute the distribution of the variable length conditional on the values of the variable age, as shown in Table 29.

I use the conditional quartiles of the variable Length/Age to compute 80 random values for each of the ranges of the variable Length. Therefore, I have 400 random values of the variable Length for each range of the variable Age.

Finally, each firm year-observation in the simulated sample is composed of 4 randomly generated numbers. The first two numbers are a duple of Age and Length, the third number is one of the random values of internal funds, and the fourth number is one of the random values of the firm's profitability factor θ_t.

Table 18 Ranges for random sample generation based on the quantiles of AGE and LENGTH/AGE

		VALUES FOR RANGES OF LENGTHMAGE					
RANGES OF AGE	*0 to 8*	0	0.65	0.85	1	1	1
	9 to 14	0	0.5	0.75	0.91	1	1
	15 to 20	0	0.44	0.68	0.85	1	1
	21 to 28	0	0.35	0.57	0.78	0.95	1
	29 to 60	0	0.3	0.47	0.63	0.86	1

I use the FOC for the demand for capital established in equation (20) and the random generated values of internal funds, age and length to calculate the firms' optimal values of capital for each year and firm in the simulated sample. I run three separated simulations in order to compare the importance of the two possible explanations for the u-shaped investment curve. In the first simulation I assume that firms face both asymmetry of information (based on the distribution of age and length of relationship) and bankruptcy costs. In the second simulation I assume that bankruptcy costs are zero and that the only financial imperfection is the asymmetry of information between firms and banks. Finally, in the last simulation I assume that the length of relationship is always equal to the age of the firm and that therefore, the only financial imperfection is the presence of bankruptcy costs.

In Figure 14 I show the sample of simulated assets assuming both asymmetry of information and bankruptcy costs. In Figure 15 I show the results assuming only asymmetry and in Figure 16 I show the sample with only bankruptcy costs. In these figures I depict the simulated value of assets versus the level of internal funds W_0 and the value of assets in logs versus the ratio of internal funds to assets, i.e. W_0/k_1. The reason to show the case of the logarithm of assets versus the ratio of internal funds to assets is to be able to validate the model specification that I will run in the empirical exercise. When using real data, a regression using both assets and internal funds in levels faces several statistical problems that could affect the results. Additionally, in reality many firms have negative internal funds and therefore, we cannot use internal funds in logs as an explanatory variable. Thus, the most feasible regression specification is the one that uses the log of assets versus the ratio of internal funds to assets.

In all the three simulations we can see a clearly non-monotonous relation between assets and internal funds. However, the results for the first and third simulations are quite similar as we can see in Figure 14 and in Figure 16 respectively. In both of them the simulated values of assets are pretty small at the lowest range of internal funds and in both cases assets are also initially decreasing on wealth. Afterwards, there is wide range of firms for which the simulated values of assets are strongly increasing on internal funds and finally, we can see that for those firms with the largest values of wealth there is no relationship between the two variables. The difference between the first and the third simulations is basically that in the simulation when only bankruptcy costs are considered the simulated values of assets at the lowest range of internal funds are much smaller and therefore, the decreasing part of the relationship is more difficult to notice.

On the contrary, the sample of firms when only asymmetry of information is considered looks quite different. In Figure 15 we can see that the simulated values of assets of those firms with the lowest levels of internal funds are very similar to the ones of those with the largest levels of internal funds. Only the firms with a middle-low level of internal funds show a smaller value of assets. In this case the relationship between assets and internal funds is clearly u-shaped for the first lower half of the sample and there is no relationship between the two variables in the second upper half of the sample.

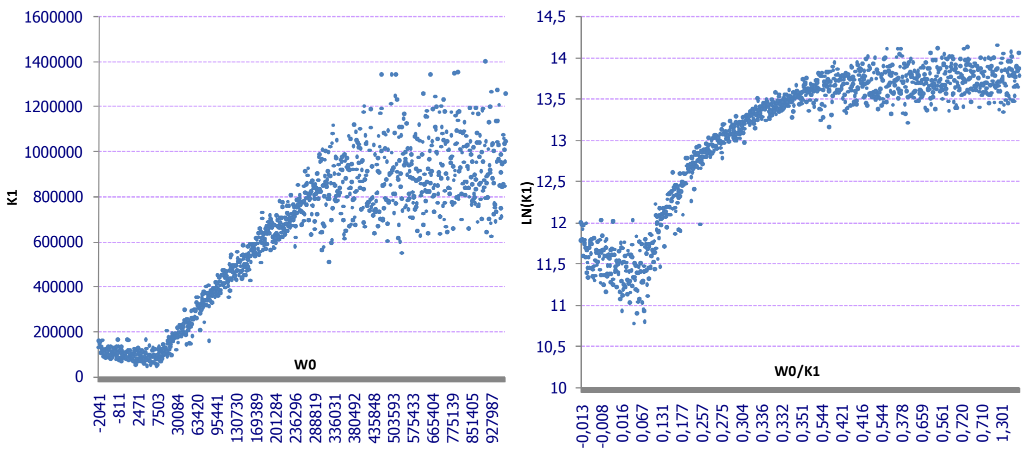


Figure 14 Simulation of firms' assets with both asymmetry of information and bankruptcy costs

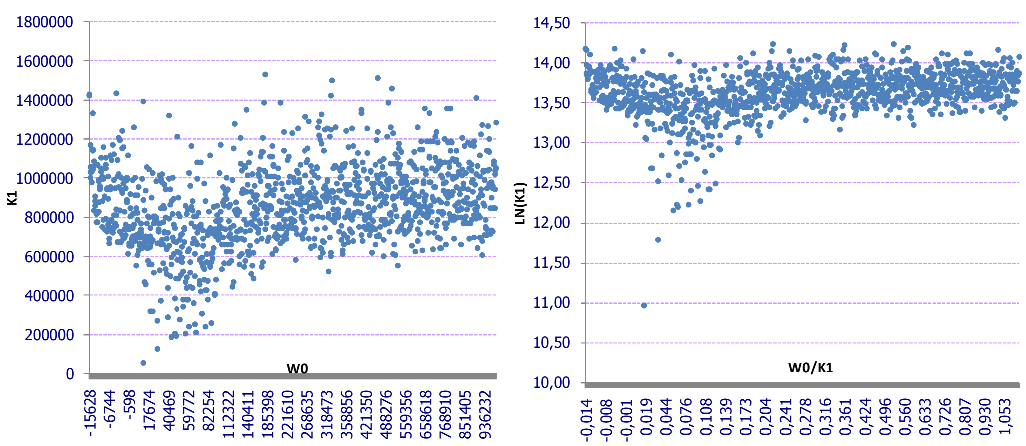


Figure 15 Simulation of firms' assets assuming only the presence of asymmetry of information

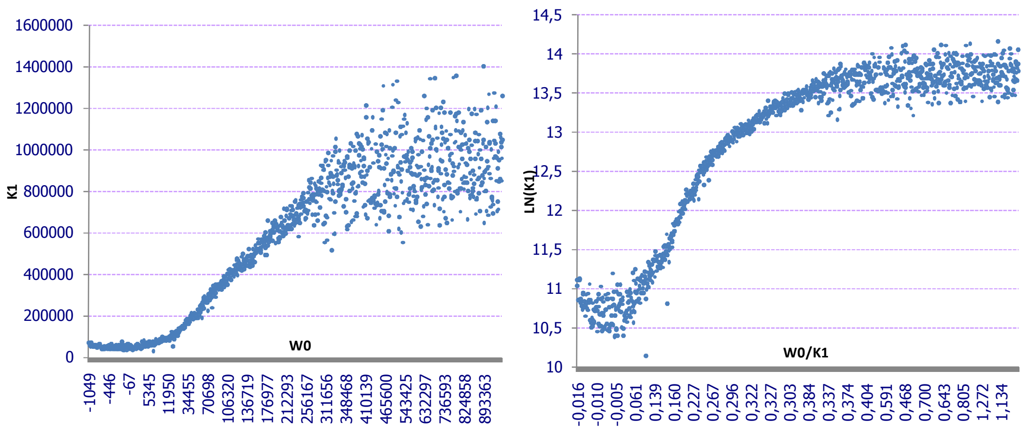


Figure 16 Simulation of firms' assets assuming only the presence bankruptcy costs

I finally run a panel regression with fixed effects with the capital and the log of capital as dependent variables. I run two different specifications of a spline regression. In the first specification the dependent variable is the level of capital and as explanatory variables I include the optimal level of capital and the financial wealth variable. The financial wealth is divided in four continuous ranges based on the four quartiles of this variable, although in some cases I use only 3 ranges of wealth. In the second specification the dependent variable is the log of assets and I replace the financial wealth by the wealth to assets ratio. In the case of the sample with only asymmetry of information I use only 3 ranges of wealth because in Figure 15 only 3 clear ranges can be distinguished where the slope of the relationship between the two variables is different. In this case I have also estimated the regression with 4 ranges and the results do not change. In the case of the two other samples (Figure 14 and Figure 16) the results are a bit sensitive to the number of ranges that I use in the spline regression, probably because the thresholds where the slope of the relationship changes are more difficult to identify statistically. I therefore show the regressions with the number of ranges that capture better what is observed in the graphs.

The results of these spline regressions for the capital in levels and in logs are shown in columns 1 and 2 of Table 19 for the simulated sample with both bankruptcy costs and asymmetry of information. The results of the regressions for the sample with only asymmetry are shown in columns 3 and 4, and the results for the sample with only bankruptcy costs are shown in columns 5 and 6. Although in the cases that the dependent variable is in levels I use the financial wealth in levels and in the cases that the dependent variable is in logs I use the wealth to assets ratio, in both cases I denote each range of these variables as "W0_range" from 1 to 4. In Table 20 I show the sensitivities of capital to wealth implied by the spline regression in each range of financial wealth.

It can be seen that in the case of the simulated sample with both asymmetry and bankruptcy costs (columns 1 and 2 in Table 19 and Table 20), the sensitivity of capital to financial wealth is first negative in the first range of observations, it becomes positive in the second range and in the next range starts decreasing and becomes zero in the third or fourth range. In both specifications, the sensitivities in each range of wealth are significant, although in the case of Assets vs. Wealth, the first negative sensitivity is only significant at the 10% level.

In the case when only asymmetry is considered (columns 3 and 4 in Table 19 and Table 20) the results are clearer. The sensitivity is negative in the first range of observations, positive in the second and near zero in the third range, with all the differences highly significant.

Finally, in the case when only bankruptcy costs are assumed (columns 5 and 6 in Table 19 and Table 20) the spline regression cannot capture the small u-shaped in the lowest range of observations. In the specification of Assets vs. Wealth the sensitivity in the first range is positive but non-significant, i.e. statistically equal to zero. In the specification with the Log of Assets vs. the wealth to assets ratio the sensitivity in the first range is negative, but again, not statistically significant.

What these results indicate is that although the theoretical model presented herein develops two alternative explanations for the existence of a u-shaped (non-monotonous) relation between capital and internal funds, an explanation based only on the presence of bankruptcy costs is not enough to find an empirical u-shaped, non-linear, non-monotonous relationship between capital and internal funds. If the empirical evidence actually confirms the existence of such relationship, the most likely explanation for such finding should be the asymmetric learning between a creditor and a firm.

Table 19 Regressions results in the simulated sample of firms

Fixed-effects (within) regression	Number of obs = 2382	Number of obs = 2382	Number of obs = 2396	Number of obs = 2396	Number of obs = 2288	Number of obs = 2288
Group variable: firm	Number of groups = 582	Number of groups = 582	Number of groups = 598	Number of groups = 598	Number of groups = 598	Number of groups = 598
	R-sq: = 0.916	R-sq: = 0.9353	R-sq: = 0.7088	R-sq: = 0.4707	R-sq: = 0.9274	R-sq: = 0.9491
	Wald chi2(4) = 12387.1	Wald chi2(5) = 15495.55	Wald chi2(4) = 3849.87	Wald chi2(4) = 1260.95	Wald chi2(4) = 15179.55	Wald chi2(5) = 14277.93
	Prob > chi2 = 0.000	Prob > chi2 = 0.000	Prob > chi2 = 0.000	Prob > chi2 = 0.000	Prob > chi2 = 0.000	Prob > chi2 = 0.000
	k1	**lnk1**	**k1 (Asymm)**	**lnk1 (Asymm)**	**k1 $(\tau>0)$**	**lnk1 $(\tau>0)$**
K1_optimal	0.4594***		0.9017***		0.4140***	
	(0.000)		(0.000)		(0.000)	
LN(K1_optimal)		1.0307***		0.9801***		0.9580***
		(0.000)		(0.000)		(0.000)
W0_range1	-2.8178*	-4.4623***	-2.4640***	-2.0431***	0.7159	-0.6568
	(0.092)	(0.000)	(0.000)	(0.000)	(0.766)	(0.447)
W0_range2	7.0424***	12.4798***	3.1216***	2.4738***	3.6206	9.2988***
	(0.005)	(0.000)	(0.000)	(0.000)	(0.138)	(0.000)
W0_range3	-3.7880***	-5.3555***	-0.6531***	-0.4031***	-3.8617***	-4.3233***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	0.0000***
W0_range4		-2.1131***				-3.5449***
		(0.000)				(0.000)
_cons	-348307***	-2.6541***	20995	0.2034	-360866***	-2.4619***
	(0.000)	(0.000)	(0.231)	(0.604)	(0.000)	(0.000)

***,**,* indicate significance at 1%, 5% and 10% respectively. P-values are in parenthesis

Table 20 Implied Sensitivities according to Spline Regressions

	k1	lnk1	k1 (Asymm)	lnk1 (Asymm)	k1 (τ>0)	lnk1 (τ>0)
Range1	**-2.818**	**-4.462**	**-2.464**	**-2.043**	**0.716**[a]	**-0.657**[a]
Range2	4.225	8.017	0.658	0.431	4.336	8.642
Range3	0.437	2.662	0.004	0.028	0.475	4.319
Range4		0.549				0.774

[a]Non-significant, i.e. statistically equal to zero

4.5 Analysis when Internal Funds are Negative and Comparison to Previous Theoretical Results

The main theoretical study that is comparable to the present one because of its similar assumptions and results is the one of Clery, Povel and Raith (2009), (CPR hereafter). The U-shaped investment curve in their study is driven by the mathematical result that $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ is always positive at positive levels of internal funds, but that it could be negative at negative levels of internal funds.

Therefore, CPR claim that for positive levels of internal funds, investment is more risky and more expensive the higher investment and debt are, but that, on the other hand, at negative levels of internal funds, investment could be less risky and less expensive the higher investment and debt are. CPR claim that given that external funds could be less expensive the more capital the firm invests when internal funds are negative, the firm can optimally choose higher levels of investment in those cases. This feature of the model by CPR is unrealistic, and in fact under the more general assumptions of my model, it is not a longer necessary condition to derive the U-shape investment function.

Therefore in this section I show that although in the present theoretical framework $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ could be sometimes negative for negative levels of wealth, the presence of a negative $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ does not drive the results presented here in any of the analyzed cases. In fact, I show that for a u-shaped investment curve to exist, we need $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ to be positive.

4.5.1 First Case: First-Best Case and Asymmetry without Bankruptcy Costs

According to the first order condition and the slope of the investment curve (given by the derivative of optimal capital with respect to internal funds), in this case the sign of $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ does not play a role:

FOC:

$$\beta \tilde{\Theta}_{1F} k_1^{*\beta-1} = (R-q)\frac{1}{\left((1-\Phi_F^{EVD})\frac{(1-\Phi_B)}{(1-\Phi_F)} + \Phi_B^{EVD}\right)}$$

Slope of the investment curve:

$$\frac{\partial k_1^*}{\partial W_0}=\frac{\frac{\partial\overline{\Theta}_1}{\partial W_0}\left(\overline{\Theta}_1-\frac{(\tilde{\Theta}_{1F}\mid S)}{(1-\Phi_F)}\right)\left(\phi_B-\phi_F\,\frac{(1-\Phi_B)}{(1-\Phi_F)}\right)\beta k_1^{\beta-1}}{-\frac{\partial V(k_1^*,\overline{\Theta}_1)}{\partial^2 k_1^*}}$$

In fact, in this case for getting a u-shaped curve we need $\frac{\partial\overline{\Theta}_1}{\partial k_1}$ to be positive. Here it is important to understand that what drives the u-shaped curve is not that the interest rate decreases at higher level of risk, but that at high levels of risk the difference on the estimated probability of default and therefore the difference in the perception of risk disappears; thus the marginal cost converges to the same one that in the first best case.

If $\frac{\partial\overline{\Theta}_1}{\partial k_1}$ were negative at the optimal, the relation between investment and internal fund would actually change from a u-shaped to an inverted s-shaped curve, since if the default point starts decreasing at some high level of risk, the default point would start approaching an intermediate level instead of a high level, and thus, the difference in the perception of risk would become larger instead of becoming smaller. Thus, the marginal cost would increase and investment would decrease.

In any of the analyzed cases the presence of negative internal funds and the existence of some ranges of investment where $\frac{\partial\overline{\Theta}_1}{\partial k_1}$ is negative is not a sufficient condition for the existence of a u-shaped curve. In order to obtain a u-shaped curve from a negative $\frac{\partial\overline{\Theta}_1}{\partial k_1}$ we would need $\frac{\partial\overline{\Theta}_1}{\partial k_1}$ to be negative at the optimal level of capital, but this is never the case.

For instance, following the proof of Proposition 1 in Appendix A. 3 it can be shown that at the first-best level of investment, $\frac{\partial\overline{\Theta}_1}{\partial k_1}$ must be positive independently of whether internal funds are positive or negative.

The explanation of this result is based on the expression that defines the default point given by the equation (A.3. 7) in Appendix A. 3:

$$\frac{\partial\overline{\Theta}_1}{\partial k_1}=\frac{(R-q(1-\tau))-\beta k_1^{\beta-1}\left(\overline{\Theta}_1(1-\Phi_B)+(\tilde{\Theta}_{1B}\mid D)\right)}{k_1^{\beta}(1-\Phi_B)}$$

The sign of the derivative of the default point with respect to the level of capital depends only on the numerator since the denominator is strictly positive. This expression has exactly the same structure as in CPR. To show that the two expressions in the two models are exactly the same, I could impose that $q=0$ and using the same notation as in CPR, I can also denote

$(1-\Phi_B)\overline{\Theta}_1 \beta k_1^{\beta-1}$ as $\int_{\overline{\Theta}_1}^{\infty} V_{k_1}(k_1, \overline{\Theta}_1)\phi_B d\Theta_1$, and $(\tilde{\Theta}_{1B} \mid D)\beta k_1^{\beta-1}$ as $\int_0^{\overline{\Theta}_1} V_{k_1}(k_1, \Theta_1)\phi_B d\Theta$, and therefore, I can express the numerator in the following way[13]:

(31)
$$R - \int_{\overline{\Theta}_1}^{\infty} V_{k_1}(k_1, \overline{\Theta}_1)\phi_B d\Theta_1 - \int_0^{\overline{\Theta}_1} V_{k_1}(k_1, \Theta_1)\phi_B d\Theta_1$$

Using any of these expressions, either the equation (A.3. 7) or equation (31) it can be proved that at the optimal level of investment, $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ must be positive independently of whether internal funds are positive or negative. Following the proof of Proposition 1 at Appendix A. 3, we can use the fact that for the expected value of the firm to be positive it is necessary that $\tilde{\Theta}_{1B} \geq \overline{\Theta}_1(1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)$.

Using L'Hospital's Rule it can also be shown that $\overline{\Theta}_1(1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D) \to \tilde{\Theta}_{1B}$ when $\overline{\Theta}_1 \to \infty$ and that $\overline{\Theta}_1(1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D) \to 0$ when $\overline{\Theta}_1 \to 0$. This is what is shown in Figure 17, where I have depicted an example of $\tilde{\Theta}_{1B}$ as "Expected Value", and $\left[\overline{\Theta}_1(1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)\right]$ as "(1-prob)*D.P.+E.V.D.".

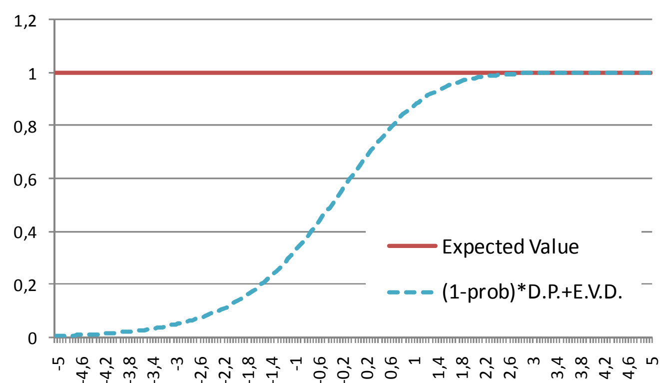


Figure 17

We can now then consider the capital level k_1^* which corresponds to the "first-best" according to bank's information, i.e. k_1^* corresponds to the level of capital such that $\tilde{\Theta}_{1B}\beta k_1^{*\beta-1} = R - q$. We can denote the default point associated to k_1^* as $\overline{\Theta}_1^*$.

At this point we have proved that $\tilde{\Theta}_{1B} \geq \overline{\Theta}_1^*(1-\Phi_B)^* + (\tilde{\Theta}_{1B} \mid D)^*$, because otherwise the value of the firm would be negative. Since the FOC tells us that at the point $\left(k_1^*, \overline{\Theta}_1^*\right)$,

[13] The expression in CPR is the following: $\int_{\underline{\theta}}^{\widehat{\theta}} F_I(I,\theta)\omega(\theta)\mathrm{d}\theta + \int_{\widehat{\theta}}^{\overline{\theta}} F_I(I,\widehat{\theta})\omega(\theta)\mathrm{d}\theta - 1$

$\tilde{\Theta}_{1B}\beta k_1^{*\beta-1} = R-q$, and since $R-q(1-\tau) > R-q$ this implies that $R-q(1-\tau) > \tilde{\Theta}_{1B}\beta k_1^{*\beta-1} = R-q > \left(\overline{\Theta}_1(1-\Phi_B)+(\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{*\beta-1}$.

Thus, $R-q(1-\tau)-\left(\overline{\Theta}_1(1-\Phi_B)+(\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{*\beta-1} > 0$ and therefore, around the point $\left(k_1^*, \overline{\Theta}_1^*\right)$ the numerator of $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ is positive and this implies that $\frac{\partial \overline{\Theta}_1}{\partial k_1}\left(k_1^*, \overline{\Theta}_1^*\right) > 0$.

What this is saying is that it does not exist an optimal level of capital at which the default point is decreasing with investment. $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ could be negative at other, lower levels of capital than the optimal, but this is not sufficient to obtain a u-shaped investment curve.

4.5.2 Second Case: Symmetry of Information but Positive Bankruptcy Costs

The case of symmetry of information but positive bankruptcy costs is more complicated mathematically than the previous one. The analysis in this case is more important, since it is the case that is closer to the one in CPR.

According to the FOC, in this case it is clear that the sign of $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ could play a role if it were negative, because a negative $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ would decrease the marginal cost of investment:

FOC:

$$\frac{\partial V(k_1^*, \overline{\Theta}_1)}{\partial k_1} = \beta \tilde{\Theta}_{1F} k_1^{*\beta-1} - (R-q) - \tau q \Phi_B - \tau q k_1 \phi_B \frac{\partial \overline{\Theta}_1}{\partial k_1} = 0$$

If we look at the derivative of the optimal level of capital with respect to internal funds, we can see that the sign of $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ also could play a role in the sign of the relationship, according to the following expression:

$$\frac{\partial k_1^*}{\partial W_0} = \frac{-\frac{\partial \overline{\Theta}_1}{\partial W_0}\tau q \phi_B \left[(1-\beta) + \frac{\partial \overline{\Theta}_1}{\partial k_1^*} k_1^* \left[\frac{(-\overline{\theta}_1 - 3\sigma_B^2/2 + \theta_B)(1-\Phi_B) + \phi_B \overline{\Theta}_1 \sigma_B^2}{\overline{\Theta}_1 \sigma_B^2 (1-\Phi_B)}\right]\right]}{-\frac{\partial V(k_1^*, \overline{\Theta}_1)}{\partial^2 k_1^*}}$$

Since the denominator is positive, $(-\frac{\partial V(k_1^*, \overline{\Theta}_1)}{\partial^2 k_1^*} > 0)$, and given that $-\frac{\partial \overline{\Theta}_1}{\partial W_0} > 0$, the relationship between the optimal level of capital and internal funds is positive whenever the term in square brackets in the numerator is positive.

Since $(1-\beta) > 0$ the term in square brackets will be strictly positive whenever the following expression is positive:

$$\frac{\partial \overline{\Theta}_1}{\partial k_1^*} k_1^* \left[\frac{-\left[\overline{\theta}_1 - (\theta_B - 3\sigma_B^2 / 2)\right](1-\Phi_B) + \phi_B \overline{\Theta}_1 \sigma_B^2}{\overline{\Theta}_1 \sigma_B^2 (1-\Phi_B)} \right]$$

It is important to remind that since $\overline{\theta}_1$ is normally distributed, it is negative at lower values than the mean and positive at higher values than the mean. Thus, the default point $\overline{\theta}_1$ is negative when internal funds are high and positive when internal funds are low.

a) *Analysis if internal funds are positive:*

In the case that W_0 is positive it has been shown that $\frac{\partial \overline{\Theta}_1}{\partial k_1^*}$ is strictly positive. In such case, we know that it is guaranteed that the relationship between investment and internal funds is positive whenever the default point is lower than the mode of the distribution, i.e. whenever $\overline{\theta}_1 \leq \theta_B - 3\sigma_B^2 / 2$, because in this case the whole numerator is positive. When the default point is positive and higher than the mode, i.e. whenever $\overline{\theta}_1 > \theta_B - 3\sigma_B^2 / 2$, there could be cases when the first term (negative) in the numerator is larger than the second term and the relationship becomes negative.

b) *Analysis if internal funds are negative:*

When internal funds are negative the previous analysis changes completely. When W_0 is negative, we know that the default point reaches its highest (positive) levels and therefore, we should have that it is higher than the mode, i.e. $\overline{\theta}_1 > \theta_B - 3\sigma_B^2 / 2$. Therefore, the term $-\left[\overline{\theta}_1 - (\theta_B - 3\sigma_B^2 / 2)\right](1-\Phi_B)$ is always negative. If the latter is sufficiently larger (negative) and thus the whole numerator is negative, i.e. if $-\left[\overline{\theta}_1 - (\theta_B - 3\sigma_B^2 / 2)\right](1-\Phi_B) + \phi_B \overline{\Theta}_1 \sigma_B^2 < 0$, then the relationship between optimal capital depends on the sign of $\frac{\partial \overline{\Theta}_1}{\partial k_1^*}$. If $\frac{\partial \overline{\Theta}_1}{\partial k_1^*}$ is negative, it would be multiplied by another negative term and therefore, the whole derivative would be positive and the relationship between optimal capital and internal funds would become immediately positive.

On the other hand, if at the optimal $\frac{\partial \overline{\Theta}_1}{\partial k_1^*}$ is positive, even though internal funds are negative, the relationship between investment and internal funds would be negative and the analysis would be exactly the same as in the case of positive internal funds.

Therefore, for the relationship between investment and internal funds to be negative when internal funds are negative, we actually need that at the optimal level of capital, the sign of $\frac{\partial \overline{\Theta}_1}{\partial k_1^*}$ is positive.

What I have shown is that when internal funds are negative, for the relationship between investment and wealth to be negative, the sign of $\frac{\partial \overline{\Theta}_1}{\partial k_1^*}$ must be positive at the optimal level of capital. Similar to the case of asymmetry of information, this does not mean that $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ could be negative at some levels of capital, but that it cannot be positive at the optimal.

4.5.3 General Analysis and Graphical Explanation

An intuitive explanation of why there are some levels of capital at which the risk is decreasing on the amount of debt is the following: If the firm is already indebted at $t = 0$, i.e. if $W_0 < 0$, it would default with probability one if it does not invest anything, because it would have zero revenues to repay the existing debt, i.e. if $W_0 < 0$ then $\Phi_B = 1$ if $k_1 = 0$.

If internal funds are negative and large enough, the firm will default for sure for a wide range of levels of capital, because at those levels, the return of investment is not enough to pay for the existing burden of debt. Therefore, there should be a minimum value of capital that a firm would need to invest to be able to repay the existing debt with a positive probability. At that minimum level, the probability of default must be close to one, and the more the firm invest the more revenues available to repay the large initial debt. This implies that the probability of default could be initially decreasing for certain small levels of capital when $W_0 < 0$. But given that the firm needs to incur into higher levels of debt to be able to increase investment, the risk must start increasing again at some point.

On the other hand, when $W_0 \geq 0$ it means that the firm does not owe any debt from a previous period or that investment does not carry any fixed cost, and thus, the only risk of default can only come from the debt that the firm acquires to finance its current investment capital.

Another key difference among the two theoretical explanations is that CPR assume that $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ is always negative when internal funds are negative. I show that this is not always the case, and that even if internal funds are negative, $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ is negative only for some levels of capital, but not for the entire domain of the profits function. Moreover, as explained before, $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ must be positive at the optimal, even if internal funds are negative. The only case in which $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ could be negative at the level of capital that maximizes profits is when profits are negative for every possible value of capital, i.e. if $E_0 V_0 \left(k_1, \overline{\Theta}_1\right) <= 0, \forall k_1$.

In other words, even if there exist some cases in which the level of risk decreases when the firm demands more debt, none of those cases corresponds to an optimal level of capital, neither the first-best level or the optimal level under asymmetry of information, positive bankruptcy costs or the latter two cases together.

I also show this argument graphically in Figure 18 for the first-best level of capital; and in Figure 19 and Figure 20 for the asymmetric case and the bankruptcy costs case respectively. The figures depict different functions. The graphs at the top show the expected value of the firm (profits), for a negative value of internal funds in each of the cases. The value of internal funds is W_0=-2000 for the first best case, W_0=-500 for the asymmetric case and W_0=-300 for the bankruptcy costs (assuming τ=0.1).

In all the cases, the expected value of the firm is negative at the points at which the revenues are not enough to repay the negative wealth and the new debt. Obviously, if the firm does not invest anything (k_1=0) the expected value of the firm is negative and the firm defaults with probability one. Therefore, the default point is decreasing for the lowest levels of capital, as it can be seen in the graphs at the bottom of each figure. However, in all of the cases we can see that although the default point is first decreasing, it is always increasing at the optimal level of capital.

The graphs at the middle of the figure show the marginal product of investment, the marginal cost in every case, and the functions $R - q(1-\tau)$ and $\left(\overline{\Theta}_1 (1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{\beta-1}$ in each of the cases as well. The function $\left(\overline{\Theta}_1 (1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{\beta-1}$ is called "marg.prod.def.point". We can see that in all the cases the graph of $\left(\overline{\Theta}_1 (1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{\beta-1}$ is first higher than $R - q(1-\tau)$ at the lowest levels of capital, and thus at those levels $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ is negative (the default point decreases). However, we can see that in all the cases $\left(\overline{\Theta}_1 (1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{\beta-1}$ becomes lower than $R - q(1-\tau)$ at a level of capital smaller than the optimal one, and thus, at the optimal, $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ is always positive (the default point increases).

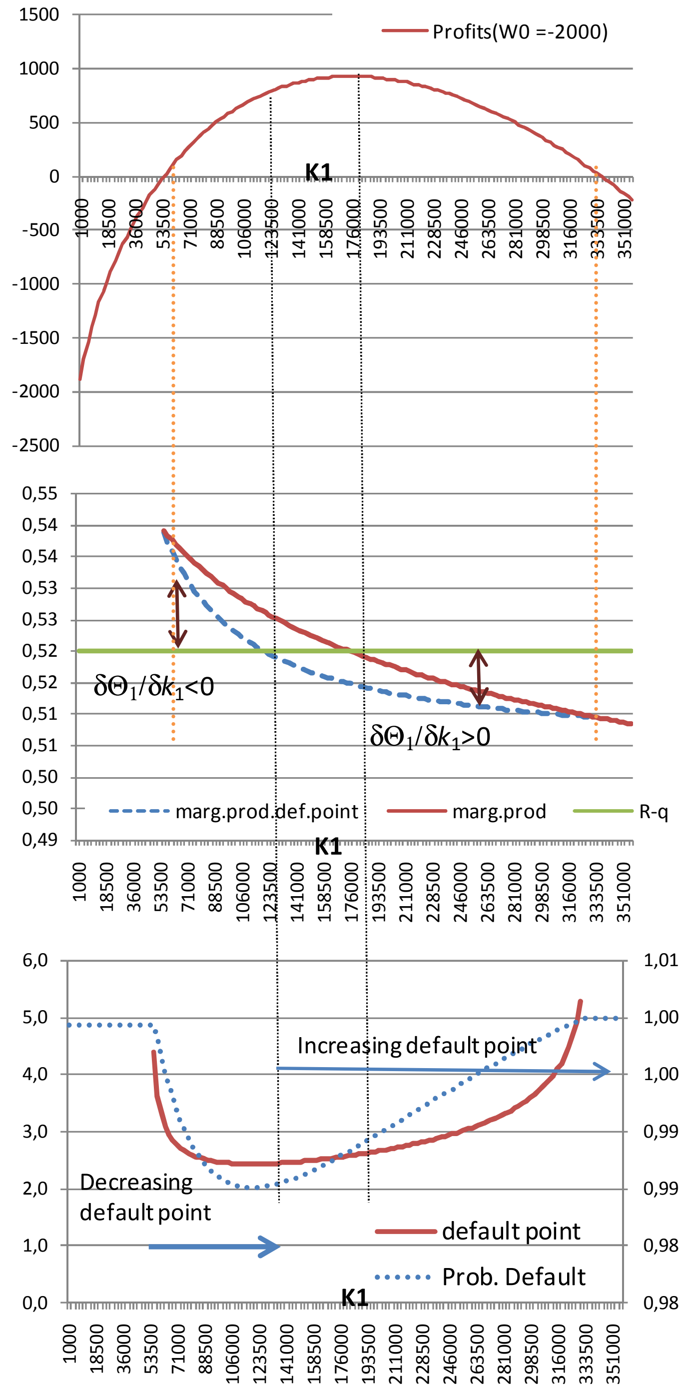


Figure 18 First best case

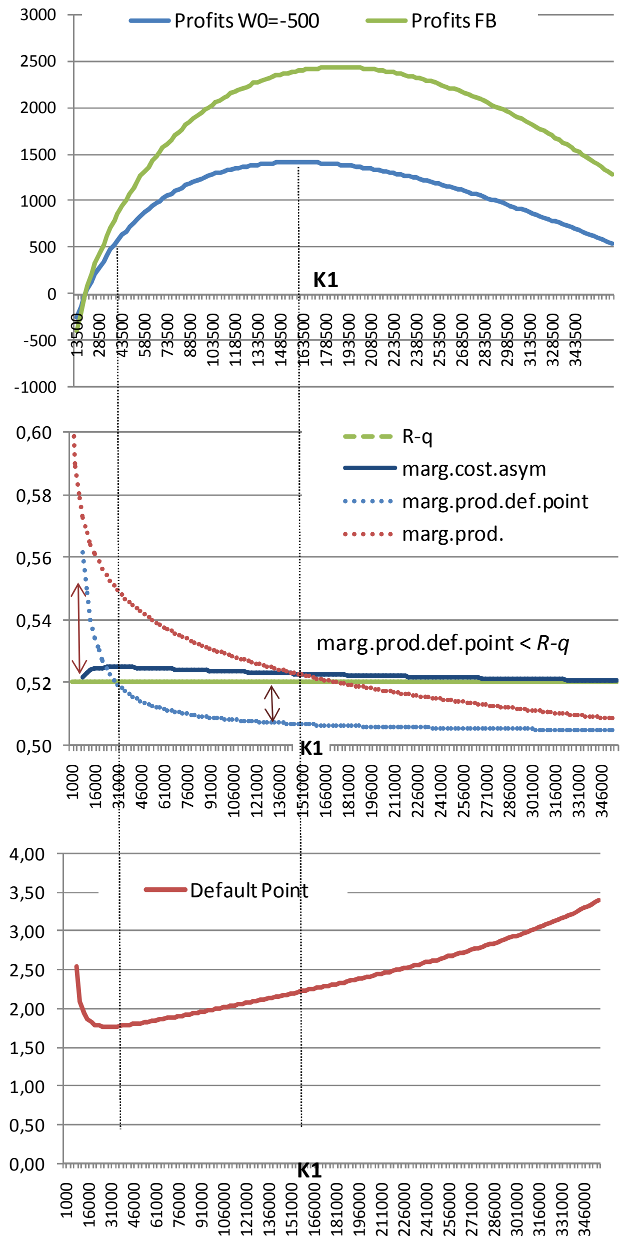


Figure 19 Asymmetric case

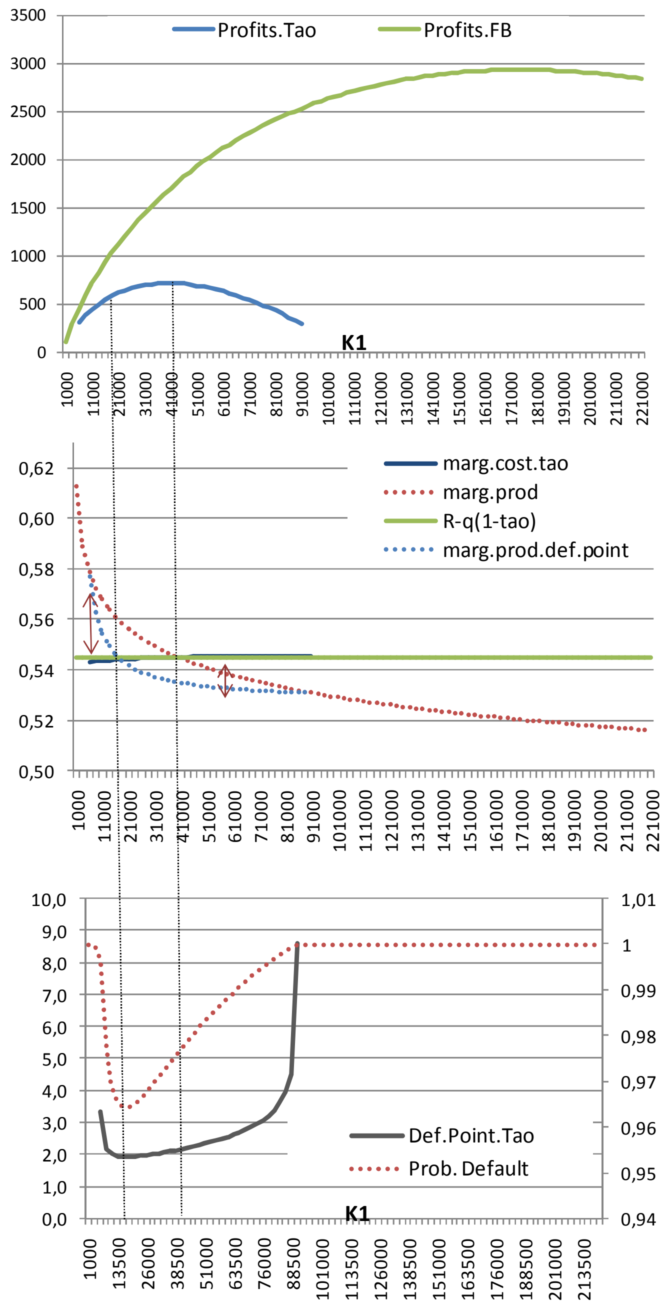


Figure 20 Bankruptcy costs (symmetric)

4.6 Empirical Analysis

4.6.1 Data and Variable Definitions

For testing the prediction of a non-linear relationship between investment and internal funds I estimate a model based on equations (26) and (29). In these equations the level of investment depends on the estimated profitability level $\ln \tilde{\Theta}_{1F}$, on the risk-free rate and the depreciation rate *(R-q)*, and on the parameter β. In the cases with financial constraints it also depends on the initial level of internal funds W_0 through its effect on the default point and therefore on the estimated default probabilities and on the conditional expected values as it can be seen in equations (26) and (29).

Since in reality the parameter β could be different for different firms, the parameter β would be captured by a firm's specific effect. The risk-free rate *R* varies through time, and thus, the estimation would require a time effect. In the equations (26) and (29) we can see that if there were no financial constraints ($\tau = 0$ and with symmetry of information) the level of capital would depend only on the log of the profitability factor $\ln \tilde{\Theta}_{1F}$, which is the forecasted profitability or productivity level of the firm.

If on the other hand, there is asymmetry and/or positive bankruptcy costs then the level of capital should depend on the level of risk, and thus on the level of internal funds. In other words, when there is asymmetry or bankruptcy costs the terms $\left((1-\Phi_F^{EVD})\frac{(1-\Phi_B)}{(1-\Phi_F)}+(\Phi_B^{EVD})\right)$ or $\tau q\left(\Phi_B+\phi_B k_1^* \frac{\partial \overline{\Theta}_1}{\partial k_1^*}\right)$ vary with the level of internal funds W_0 (or in general with any variable besides the firm's profitability that affects the probability of default and the conditional expected values of the firm in the cases of default and success.

I test the prediction for different definitions of internal funds, which are financial wealth, cash or equity. Therefore, the basic equation to estimate would be:

$$\textbf{(32)} \qquad \ln(FIXDCAP)_{i,t} = \beta_1(PROFIT)_{i,t-1} + \beta_2(WEALTH)_{i,t-1} + \alpha_i + \delta_t + \upsilon_{it}$$

Where lnFIXEDCAP denotes the natural logarithm of fixed assets, *WEALTH* denotes financial wealth divided by assets. PROFIT denotes a moving average of a proxy for the profitability factor $\ln \tilde{\Theta}_{1F}$. Financial wealth is defined as the sum of the variables cash in hand, accounts receivables and inventory minus the variable accounts payable. δ_t denotes a time specific effect that captures the changes in the risk-free rate and α_i denotes a firm specific effect.

Since the dependent variable in equation (32) is the log of fixed capital, ideally I should use the financial wealth variable also in logs. However, I cannot use the log of wealth because nearly one third of the firms have negative values of wealth and therefore, I would lose an important number of observations. This is why I decided to use the wealth to assets ratio as independent variable as a proxy of the firm's level of risk. Simulation results confirm the validity of approximating equations (26) and (29) by regressing the log of capital on the wealth to assets ratio, as it is done in equation (32).

Since the vast majority of firms in the sample are non-quoted, I cannot compute the Tobin's Q to control for the availability of positive net-present-value projects. Therefore I proxy for the expected profitability level of the firm using a compound index of the sales to assets ratio together with the profits to assets ratio, trying to account for different aspects of profitability. The compound index is estimated through a Principal Component Analysis of the Sales/Assets and Profits/Assets ratios. The correlation of the Sales/Assets and the Profits/Assets ratios with the compound index of profitability is 0.71 and 0.72 respectively. I choose to use a combined index of sales and profits because in the theoretical model I assume that the profitability factor Θ_t may vary because of random variations in productivity (profits), but also because of random variations in the demand (sales). Since in the theoretical model the estimated profitability factor $\tilde{\Theta}_{1F}$ is the average of the whole set of observations at each point in time, I use a moving average of the compound index for each firm. The moving average is different at each point in time, since it is computed as the average between time 0 and time *t*, at each *t*. Although not reported here, I also estimate a second version in which instead of controlling for the average of the profitability index I use the estimated productivity shock from balance sheet data in a similar way to Caggese (2007).

One potential problem with the basic specification in equation (32) is that in reality capital is subject to convex adjustment costs and irreversibility. However such costs are of uncertain nature, and they are likely to vary across firms. Given this problem, equation (32) is possibly misspecified. However, since the theoretical model in which equation (32) is based does not include adjustment costs and/or the lagged values of fixed capital, I first estimate the model without lagged values using a simple fixed-effect (within) estimator methodology. Nonetheless, in order to take into account these possible misspecification problems, I then estimate the model with lagged capital included as an explanatory variable. Therefore, in order to have a consistent estimator, the parameters are estimated using a dynamic panel methodology and specifically applying the General Method of Moment (GMM) on the equation in first differences suggested by Arellano and Bond (1991). Additionally, I also estimate a more general flexible accelerator model replacing the log of fixed capital with the investment rate of fixed capital as dependent variable, and including the lagged value of investment and the lagged square value of investment, in order to account for the nonlinearities implied by the convex adjustment costs and irreversibility of investment.

I use a subsample of the first 100,000 firms in alphabetical order from the SABI database. The whole SABI database contains in total about 811,000 firms (after eliminating firms from the financial sector). I construct the dataset using the annual information from 1992 to 2007 period. The final sample consists of an unbalanced panel with a maximum of 347,463 firm-observations with 65,176 firms, although the sample changes depending on the different specifications.

4.6.2 Spline Regression, Basic Specification and Robustnes checks

Since the most adequate specification to account for non-linearities in the data is the spline regression, I modify equation (32) to capture the sensitivity of fixed capital to seven different ranges of the wealth to assests ratio.

$$\ln(FIXDCAP)_{i,t} = \beta_1 (PROFIT)_{i,t-1} + \\ + \beta_2 (WEALTH)_{i,t-1} + \beta_3 (WEALTH*(D_1))_{i,t-1} + ... + \beta_8 (WEALTH*(D_6))_{i,t-1} + \delta_t + \upsilon_{i,t}$$

Therefore, the spline regression is based on a division of the sample in 7 continuous ranges of observations according to the following values given by the distribution of the variable $WEALTH_{i,t-1}$, with the first range represents 15% of the observations and each one of the following ranges represent the following 15% of the observations, with the exception of the highest range that represents the highest 10%. The ranges are like the following:

Range 1: If $WEALTH_{i,t-1} < -0.3112$

Range 2: $-0.3113 < WEALTH_{i,t-1} < -0.048$

Range 3: If $-0.048 < WEALTH_{i,t-1} < 0.071$

Range 4: If $0.071 < WEALTH_{i,t-1} < 0.200$

Range 5: If $0.200 < WEALTH_{i,t-1} < 0.362$

Range 6: If $0.362 < WEALTH_{i,t-1} < 0.619$

Range 7: If $WEALTH_{i,t-1} > 0.619$

As explained before, I first estimate the basic specification without including the lagged value of fixed assets as an explanatory variable and using a fixed effect (within) estimator. The results of this specification are shown in column 1 of Table 21. However, we can see that in this case the signs of the explanatory variables are exactly the opposite of the ones expected in theory. More specifically, the correlation of fixed assets with the profitability factor is negative and the sensitivities to wealth are exactly the opposite of the expected u-shaped function. Since we know that this first specification is biased and inconsistent, I then estimate the equation including the lagged value of fixed assets and estimating the equation through a dynamic panel (GMM). In this second specification shown in column 2, the signs of the coefficients are as expected, highly significant and they indicate a clear non-monotonic relation between fixed capital and financial wealth. The first sensitivity is negative as expected; it becomes positive in the second range of wealth, and stays positive afterwards.

As a first robustness check, I also estimate the basic specification but using as dependent variable the investment to assets ratio. I first estimate the equation without controlling for lagged values of investment, although such specification should also be biased and inconsistent; the results are shown in column 3. In this case, the sign of the profitability factor is positive, but non-significant, the same as almost all the rest of explanatory variables. Again, this is due to the misspecification of the equation due to the presence of adjustment costs and other factors. I thus estimate a flexible accelerator model, including the lagged value of investment and the lagged square value of investment rate in order to account for the nonlinearities implied by convex adjustment costs. This equation is also estimated through a dynamic panel (GMM) estimator. In this case the results are again as expected, because I find that the first sensitivity of investment to financial wealth is negative, it becomes positive in the second range and decreased (staying positive) at the highest ranges of wealth.

Summarizing, when the equation is estimated consistently through a dynamic panel and taking into account the correlation of assets to its lagged values, the empirical evidence confirms the presence of a non-monotonic, u-shaped function of investment.

Table 21 Basic Specification and Robustness checks. Methodology and Specification

	Number of obs = 366819	Number of obs = 298246	Number of obs = 366819	Number of obs = 258536
	Number of groups = 74015	Number of groups = 61621	Number of groups = 74015	Number of groups = 50637
	R-squared = 0.1522	Wald chi2(19) = 7066.19	R-squared = 0.004	Wald chi2(19) = 1680000
	Prob > F = 0.000	Prob > chi2 = 0.000	Prob > F = 0.000	Prob > chi2 = 0.000
	LNFXDASSETS (FIXED EFFECTS)	**LNFXDASSETS (GMM)**	**INVESTM/ASSETS (FIXED EFFECTS)**	**INVESTM/ASSETS (GMM)**
LNFXDASSETS (t-1)		0.5552***		
		(0.000)		
INV/ASSETS (t-1)				0.2089***
				(0.001)
INV/ASSETS^2 (t-1)				0.0025***
				(0.000)
PROFITABILITY FACT (t-1)	-2.7153***	4.6464***	0.6398	6.9328***
	(0.000)	(0.000)	(0.485)	(0.000)
WEALTH_1	0.3601***	-0.0759***	0.0119	-0.0612*
	(0.000)	(0.000)	(0.697)	(0.082)
WEALTH_2	-0.5895***	0.5237***	0.1479*	0.5392***
	(0.000)	(0.000)	(0.074)	(0.000)
WEALTH_3	-0.7641***	0.7032***	0.1080	-0.0670
	(0.000)	(0.000)	(0.455)	(0.759)
WEALTH_4	0.0889	-0.1707*	-0.0289	0.1767
	(0.420)	(0.058)	(0.745)	(0.133)
WEALTH_5	-0.0680	-0.0471	-0.0483	-0.1451***
	(0.540)	(0.614)	(0.313)	(0.004)
WEALTH_6	-0.2815***	0.1095	0.0395	0.0238
	(0.007)	(0.255)	(0.249)	(0.613)
WEALTH_7	-0.3537**	0.1577	0.0024	-0.2322***
	(0.032)	(0.342)	(0.968)	(0.000)
_cons	11.5452***	4.8748***	-0.0365**	-0.1278***
	(0.000)	(0.000)	(0.045)	(0.000)
Sargan chi2		60.421		58.068
		(0.198)		(0.223)
Second Order z		1.389		1.452
		(0.165)		(0.146)

Furthermore, as an additional robustness check I also estimate the previous spline regressions using variable capital instead of fixed capital as in Caggese (2007), in order to show that the results are robust to the objection that the equation is misspecified because of convex adjustment costs of capital. The results of this estimation are shown in Table 22. In the first column of this table I show that since variable assets are also highly correlated with its lagged value, a fixed-effect (within) regression that does not include the lagged value of variable assets also has a similar problem that in the case of fixed assets. When I include the lagged value of variable assets and estimate the regression through a dynamic panel (GMM) as it is done in column 2, I again obtained the same result (u-shaped) that in the fixed assets case. Furthermore, I also estimate a regression with the investment to assets ratio as dependent variable (using variable assets) and using a fixed-effects estimator. Differently to the case of fixed assets, I obtain the theoretically expected result without needing to use a specification that accounts for convex adjustment costs.

Table 22 Robustness Checks. Spline Regressions with Variable Assets

	Number of obs = 374766	Number of obs = 287860	Number of obs = 375020
	Number of groups = 75766	Number of groups = 60199	Number of groups = 75804
	R-squared = 0.1811	Wald chi2(15) = 13910.15	R-squared = 0.005
	Prob > F = 0.000	Prob > chi2 = 0.000	Prob > chi2 = 0.000
	LNVARASSETS (FIXED EFFECTS)	**LNVARASSETS (GMM)**	**INV_VARASSETS (FIXED EFFECTS)**
LNVARASSETS(t-1)		0.4364***	
		(0.000)	
PROFITABILITY FACT (t-1)	-0.2822***	0.6418*	0.1534*
	(0.008)	(0.095)	(0.084)
WEALTH_1	0.3171***	-0.6019***	-0.2620***
	(0.000)	(0.000)	(0.004)
WEALTH_2	0.1138***	0.3413***	1.0284**
	(0.000)	(0.000)	(0.042)
WEALTH_3	-0.5173***	0.3563***	-0.8835
	(0.000)	(0.000)	(0.190)
_cons	5.8641***	6.8741***	-0.1434*
	(0.000)	(0.000)	(0.060)
Sargan chi2		63.851	
		(0.208)	
Second Order z		1.899	
		(0.365)	

Finally, as a further way of checking the robustness of the results, I have also performed the same spline regressions replacing the wealth ratio by the cash-flow and equity to assets ratios. The results using cash flow ratio are shown in Table 23. Although the results using equity are not reported, all the results also hold in that case. Similar to the spline specification using financial wealth, I have constructed 4 ranges according to statistical distribution of the variable cash-flow in the SABI sample. In the first column of Table 23 we can see that the results are pretty similar to the ones using financial wealth. I also run the same robustness tests as with financial wealth. Therefore, in column 2 I show the results using Variable assets instead of Fixed Assets. Similarly, in column 3 I estimate the same equation but using Investment/Assets as dependent variable and using Fixed Effects instead of Dynamic Panel-GMM. In column 3 I use a flexible accelerator specification and a Dynamic Panel method. Finally, in column 5 I use the Investment/Assets as dependent variable, but using Variable Assets instead of Fixed Assets, and Fixed Effects instead of Dynamic Panel-GMM.

In all the cases the results obtained with the wealth to assets ratio are still obtained when using the cash flow ratio.

Table 23.

	Number of obs = 331515	Number of obs = 340858	Number of obs = 409828	Number of obs = 270678	Number of obs = 418744
	Number of groups = 62312	Number of groups = 64066	Number of groups = 75694	Number of groups = 54746	Number of groups = 77456
	Wald chi2(16) = 9031.5	Wald chi2(16) = 21414.11	R-squared = 0.006	Wald chi2(16) = 1148.46	R-squared = 0.0438
	Prob > chi2 = 0.000	Prob > chi2 = 0.000	Prob > F = 0.000	Prob > chi2 = 0.000	Prob > F = 0.000
	LN_FXDASSETS (GMM)	**LN_VARASSETS (GMM)**	**INV_ASSETS (FXD.ASSETS) (FXD EFFECTS)**	**INV_ASSETS (FXD.ASSETS) (GMM)**	**INV_ASSETS (VAR.ASSETS) (FXD EFFECTS)**
LNFXDASSETS (t-1)	0.5168***				
	(0.000)				
LNVARASSETS(t-1)		0.4087***			
		(0.000)			
INV_ASSETS(t-1)				0.0953	
				(0.228)	
INV_ASSETS^2(t-1)				0.0009	
				(0.245)	
PROFITABILITY FACT (t-1)	5.9459***	2.9618**	1.8668***	5.6878***	24.7681***
	(0.000)	(0.028)	(0.000)	(0.000)	(0.002)
CASH_1	-0.1417***	-0.3659***	-0.0617***	-0.1581***	-0.3591***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CASH_2	0.4741***	1.0283***	0.0810**	0.3155***	0.8045**
	(0.000)	(0.000)	(0.031)	(0.000)	(0.044)
CASH_3	0.0658	-0.1108	0.0893*	0.1420***	-0.3921
	(0.501)	(0.184)	(0.074)	(0.006)	(0.135)
CASH_4	-0.0329	-0.5522***	0.0803*	-0.0619	-0.3664***
	(0.662)	(0.000)	(0.069)	(0.168)	(0.001)
_cons	5.5469***	7.2133***	0.0296***	0.0339***	0.0667
	(0.000)	(0.000)	(0.000)	(0.000)	(0.122)
Sargan chi2	59.791			61.190	
	(0.214)			(0.179)	
Second Order z	1.2287			1.2123	
	(0.219)			(0.225)	

4.6.3 Comparison Between Younger and Older Firms

As a way to show that the presence of a non-linear relationship between capital and internal funds is due to the presence of asymmetry of information, I run the basic specification for firms of different ages, since it can be expected that younger firms are the ones more likely to face higher asymmetries of information and have a more uncertain distribution of revenues. Therefore, I first use those firms with ages between 3 and 10 years old and I do not include firms with ages between 0 and 2 years old because the difference between length and age for these firms cannot be too high. In other words, I try to pick firms that are not only young, but also with ages much higher than the length of the relationship[14].

In column 2 of Table 24 I show the results of the basic regression specification (including lagged fixed capital) using only the youngest firms. In column 3 of Table 24 I show the results of the same spline regression using only the sample of oldest firms, the ones older than 10 years old. We can see in column 2 that the first range shows a smaller (more negative) coefficient and that the following two ranges show higher coefficients than those for the same ranges in column 1.

[14] Ideally, I should run a separate regression for firms with lengths of relationship shorter than its age, but the SABI database does not include information on the length of relationship of firms with their creditors.

Table 24

	Number of obs = 314269	Number of obs = 67658	Number of obs = 62104
	Number of groups = 61231	Number of groups = 33978	Number of groups = 30698
	Wald chi2(20) = 8160.09	Wald chi2(20) = 3043.01	Wald chi2(20) = 810.24
	Prob > chi2 = 0.000	Prob > chi2 = 0.000	Prob > chi2 = 0.000
	LNFXDASSETS (ALL FIRMS)	**LNFXDASSETS (YOUNG)**	**LNFXDASSETS (OLD)**
LNFXDASSETS (t-1)	0.5552***	0.5703***	0.6306***
	(0.000)	(0.000)	(0.000)
PROFITABILITY FACT (t-1)	4.6464***	4.5262***	8.3762**
	(0.000)	(0.000)	(0.033)
WEALTH_1	-0.0759***	-0.0979***	-0.0585
	(0.000)	(0.000)	(0.101)
WEALTH_2	0.5237***	0.6085***	0.3994***
	(0.000)	(0.000)	(0.000)
WEALTH_3	0.7032***	0.7499***	0.7850***
	(0.000)	(0.000)	(0.000)
WEALTH_4	-0.1707*	-0.2386*	-0.0229
	(0.058)	(0.095)	(0.838)
WEALTH_5	-0.0471	-0.1532	0.0523
	(0.614)	(0.327)	(0.642)
WEALTH_6	0.1095	0.0470	0.1913
	(0.255)	(0.769)	(0.114)
WEALTH_7	0.1577	0.1709	0.2277
	(0.342)	(0.501)	(0.324)
_cons	4.8748***	4.5411***	4.1960***
	(0.000)	(0.000)	(0.000)
Sargan chi2	62.698	62.509	63.622
	(0.147)	(0.151)	(0.130)
Second Order z	1.442	1.361	1.408
	(0.149)	(0.174)	(0.159)

The coefficients estimated in this spline regression imply that the coefficients of the lagged wealth on investment for the three different samples, whole sample, younger and older firms are the following:

Table 25 Estimated Elasticities of the log of capital to financial wealth for the different ranges according to the spline regression

	ALL FIRMS	**YOUNGER**	**OLDER**
Range1	**-0.076**	**-0.098**	**-0.059**
Range2	0.448	0.511	0.341
Range3	1.151	1.260	1.126
Range4	0.980	1.022	1.126
Range5	0.980	1.022	1.126
Range6	0.980	1.022	1.126
Range7	0.980	1.022	1.126

It is clear in Table 25 that the relationship between investment and wealth is negative for the lowest range of wealth, and positive for the other 6 higher ranges of wealth for the three samples of firms. However, the sign is lower (more negative) for the younger firms and higher (less negative) for the older firms. This result clearly supports the theoretical prediction of the model. Moreover, it is clear that although the second range of firms are all firms with negative values of internal funds, their estimated sensibility to internal funds is positive, showing that what drives the results is not the presence of negative values of wealth, but the riskiness of the ones with the lower values. This clearly supports the theoretical predictions of the model in favor of other possible explanations.

4.6.4 Polynomial Specification

Since the theoretical model predicts that the relationship between investment and internal funds is not linear, I also run a polynomial specification adding the square of the wealth ratio.

$$\ln(FIXDCAP)_{i,t} = \beta_1 \ln(FIXDCAP)_{i,t-1} + \beta_2 (WEALTH)_{i,t-1} + \beta_3 (WEALTH^2)_{i,t-1} + \beta_4 (PROFIT)_{i,t-1}$$

It can be seen in Table 26 that the coefficients on the wealth terms are all significant. The coefficient on the linear term of wealth is positive and significant and the coefficient on the square term is also positive and significant. The sign of these coefficients imply that the relationship between investment and wealth is non-monotonic and it resembles a u-shaped function. The same result is found if instead of financial wealth I use cash flow ratio as it can be seen in columns 2 and 3 of Table 26. In both cases the linear and quadratic terms are both positive and significant implying a non-linear u-shaped relationship.

Table 26.

	Number of obs = 314269	Number of obs = 3311515
	Number of groups = 61231	Number of groups = 62312
	Wald chi2(15) = 8136.87	Wald chi2(15) = 8712.51
	Prob > chi2 = 0.000	Prob > chi2 = 0.000
	LNFXDASSETS	**LNFXDASSETS**
LNFXDASSETS (t-1)	0.5761***	0.5987***
	(0.000)	(0.000)
PROFITABILITY FACT (t-1)	4.3279***	6.6739***
	(0.000)	(0.000)
WEALTH (t-1)	0.4479***	
	(0.000)	
WEALTH^2 (t-1)	0.2313***	
	(0.000)	
CASH (t-1)		0.1463***
		(0.000)
CASH^2 (t-1)		0.3687***
		(0.000)
_cons	4.6807***	5.2964***
	(0.000)	(0.000)
Sargan chi2	63.296	59.531
	(0.136)	(0.221)
Second Order z	1.377	1.462
	(0.169)	(0.144)

4.6.5 Split Sample Regression

In Table 27 I show the results of the basic linear specification but splitting the sample according to the values of the wealth to assets ratio. In column 1 I show the results using the whole sample of firms. In column 2 I show the results of using only the observations with values of financial wealth smaller than the lowest quintile. Finally, in column 3 I show the results using only the sample of observations with values of lagged wealth higher than the lowest quintile.

If I estimate the regression for the whole sample, the estimated coefficient for the wealth factor is positive and significant. If I estimate the regression for the sample with only negative values of wealth, the estimated coefficient is negative (-0.0247), although non significant. On the other hand if I estimate the model for the sample with only positive values of wealth, the estimated coefficient is positive and significant and higher than in the whole sample estimation (0.5485 vs. 0.4673).

Table 27

	Number of obs = 331515	Number of obs = 28024	Number of obs = 309247	Number of obs = 347463	Number of obs = 42676	Number of obs = 304787
	Number of groups = 62312	Number of groups = 16880	Number of groups = 60002	Number of groups = 65176	Number of groups = 17474	Number of groups = 58045
	Wald chi2(14) = 8726.55	Wald chi2(14) = 64.68	Wald chi2(14) = 8830.34	Wald chi2(14) = 8770.3	Wald chi2(14) = 854.65	Wald chi2(14) = 9311.75
	Prob > chi2 = 0.000	Prob > chi2 = 0.000	Prob > chi2 = 0.000	Prob > chi2 = 0.000	Prob > chi2 = 0.000	Prob > chi2 = 0.000
	LNFXDASSETS	**LNFXDASSETS**	**LNFXDASSETS**	**LNFXDASSETS**	**LNFXDASSETS**	**LNFXDASSETS**
LNFXDASSETS (t-1)	0.5864***	0.3575***	0.5821***	0.5975***	0.4866***	0.5513***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
PROFITABILITY FACT (t-1)	4.6290***	4.4907***	4.0107***	7.1372***	7.5697***	6.6720***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
WEALTH (t-1)	0.4673***	-0.0247	0.5485***			
	(0.000)	(0.209)	(0.000)			
CASH (t-1)				0.0492***	-0.0462***	0.1218***
				(0.000)	(0.050)	(0.000)
_cons	5.3405***	5.7313***	5.2782***	5.3405***	5.7313***	5.2782***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sargan chi2	62.9350	60.6216	64.0692	63.771	64.194	59.334
	(0.142)	(0.193)	(0.122)	(0.127)	(0.120)	(0.226)
Second Order z	1.2653	1.2671	1.3766	1.364	1.387	1.325
	(0.206)	(0.205)	(0.169)	(0.173)	(0.165)	(0.185)

If I repeat the exercise using cash flow as proxy for internal funds, the results are even clearer. If I estimate the regression for the whole sample, the estimated coefficient for cash is positive and significant (column 4). If I estimate the regression using the sample with the smallest values of cash flow, the estimated coefficient is negative and statistically significant (column 5). Finally, if I estimate the model for the sample with only positive values of cash or equity, the estimated coefficient is positive, significant and much higher than in the whole sample estimation (columns 6).

4.7 Conclusions

The main added value of this chapter is the development of a new investment model with financial constraints and Bayesian learning that provides new predictions about the investment behavior of firms when creditors have a different perception of risk than the firm due to asymmetric learning, or when credit is too expensive due to bankruptcy costs. I apply this framework to investigate theoretically and empirically the relationship between firms' investment and internal funds in the presence of limited information, learning and bankruptcy costs. The theoretical predictions of the model can help to rationalise the results of recent empirical studies that have provided evidence that firms with lower levels of internal funds, that are near default, or that have negative levels or internal funds show lower or negative sensitivities to them, such as in Allayannis and Mozumdar (2004), Lyandres (2007), Cleary, Povel and Raith (2007) and Hovakimian (2009). The theoretical predictions of the model also contribute to the old debate about the usefulness of investment cash-flow sensitivities as a measure of financial constraints.

Importantly, the theoretical model provides a new framework for analyzing the possibility of asymmetric learning between a bank and a firm and its effect on firm's investment decisions and other credit related issues. This new framework proves to be useful for studying other related topics such as the evolution of interest rates during the life cycle of firms and through the duration of a bank-firm relationship, a topic that is studied in chapter 5.

The new results provided by this chapter can be summarized as follows:

i) It shows that a U-shape investment function arises naturally in the presence of asymmetric learning of banks and firms about the firms' riskiness, even in the absence of bankruptcy costs.

ii) It extends Cleary *et al* (2007) model in order to clarify under what conditions bankruptcy costs are indeed sufficient for generating the U-shape function under more general distributional assumptions.

iii) It calibrates with realistic data both the model with asymmetric learning and the model with bankruptcy costs, to show that the U-shape function is more likely to arise if asymmetric learning is present.

iv) It estimates the U-shape function empirically with firm level data, to provide further evidence that asymmetric learning is the most likely explanation of it.

The theoretical model shows that investment is non-monotonically related to internal funds in a u-shaped fashion and define two possible explanations for the u-shape: the asymmetric learning and the non-monotonicity of the density function with respect to internal funds and its interaction with the expected bankruptcy costs.

On the empirical side, this chapter confirms the presence of a negative relation between internal funds and investment for firms with the lowest levels of internal funds of firms and presents new evidence of a non-monotonic relation between investment and internal funds. According to the predictions of the theoretical model, I find that young firms, which are the ones more likely to face stronger asymmetries of information and higher uncertainty in their distribution of profits, show a clearer u-shaped relation between investment and internal funds. Therefore, both the simulation results and the empirical evidence point in the direction of the asymmetric learning explanation. Importantly, I find that not all the firms with negative values of internal funds show a negative relation between investment and the different measures of internal funds.

The main shortcomings of the theoretical model are useful to identify the most interesting extensions to it. Firstly, the model is not fully dynamic and therefore, it cannot take into account the dynamics in the firms' decision about how much financial wealth to hold and how much dividends to pay each period. Additionally, the model does not account for adjustment costs or irreversibility of investment.

Additionally, an interesting topic that derives from the model simulations is the possibility that the investment u-shaped curve changes under different macroeconomic conditions. One clear result of the simulation is that although the u-shaped exists due to both bankruptcy costs and asymmetric learning, the negative part of the u-shaped is more difficult to be observed in reality when bankruptcy costs are present. In the simulations, I use a constant value of the proportion of bankruptcy costs with respect to firms' assets, but in reality, bankruptcy costs could be much higher (in proportion) when macroeconomic conditions are worst and the market value of assets is lower. On the contrary, when the economy is in a boom, bankruptcy costs as a proportion of the market value of assets could become negligible. If this is so, the investment-internal funds relation may look more "U-shaped" in the latter case than in the former one.

CHAPTER 5

5. INVESTMENT WITH LEARNING AND FINANCIAL CONSTRAINTS: THE EFFECTS OF RELATIONSHIP LENDING ON THE COST OF CREDIT

ABSTRACT

The chapter studies the effects of relationship lending on the evolution of interest rates during the life cycle of firms. I find that when both banks and firms need to learn about firms' true profitability and to predict their probability of default, the banks' learning process may have two different effects on the evolution of interest rates through the course of a relationship. Relationship lending affects interest rates via a reduction in the uncertainty of the estimation of the probability that the firm defaults, and a reduction in the possible disagreement between the bank and the firm when there is asymmetry of information. The reduction in uncertainty always induces a reduction of interest rates, but the reduction of the asymmetry could have opposite effects depending on the firm's level of risk. This latter effect works via changes in the demand for external capital. The chapter tests and confirms the model suggestions using survey data from the US (NSSBF 2003). The findings complement mixed results from previous empirical studies on the relation of interest rates with respect to firms' age and the length of a relationship.

5.1 Introduction

The relationship of a firm with a creditor is crucial for analyzing the cost and the availability of credit of small and young firms, since it allows banks to learn about the firms they are lending to.

The banks' learning process is important because information and learning affect the perception of the risk of an investment decision. When there is little information, and the bank is uncertain about the intrinsic profitability of the firm, the business risk perceived by the bank could be higher than the risk implied by the random nature of demand, costs or productivity factors. Following this idea, we intuitively would expect that, as a bank learns more about a firm through relationship lending, the loans' interest rates would decrease[15] through the course of a relationship. If we think that through relationship lending the banks acquire more information about a firm, we would expect that this translates into a bank's lower perception of risk and therefore, into lower interest rates.

Using a model of firm investment with financial constraints and Bayesian learning I find that the process of banks' learning about a firm, represented by the length of a relationship with a firm, may have opposite effects on the interest rate paid by the firms when both banks and firms need to learn about firms' quality. In the model, firms and banks need to use their history of information about the firm to learn and predict the firm's future profits and its probability of default.

The total effect of a longer relationship lending on interest rates is the sum of two different effects that could have different signs: a reduction in uncertainty and a reduction in the disagreements between the firm and the bank when the bank has less information than the firm about the firm's intrinsic profitability. The reduction in the uncertainty always generates a reduction in the interest rate via a reduction in the variance of prediction and on the estimated probability of default. The reduction in the asymmetry of information and the possible disagreements between the firm and bank has an effect on the loan's interest rate via a change in the firm's demand for capital and more specifically, on the demand for external funds and leverage. The latter effect is higher for firms with relatively low levels of risk and is smaller for firms with high levels of risk.

The evolution of interest rates through the course of a relationship depends on which effect dominates. For those firms for which the effect on the demand for capital is small (the high risk firms) we should observe that interest rates decrease or do not change as the length of relationship increases. On the contrary, for those firms for which the reduction in the demand for capital is high enough (low and intermediate risk firms) we should observe that interest rates are initially very low (given the highly constrained demand of capital) and that they increase as asymmetry disappears and the firms demand more external funds at higher interest rates.

Furthermore, I find that this result is novel in the sense that an increase in interest rate through the course of a relationship does not come from possible lock in problems that are claimed to

[15] If we leave aside possible hold up or lock in problems generated by such relationship.

generate monopoly power for the bank through the information acquired in the course of a relationship.

I test these implications using the National Survey of Small Business Finance (NSSBF 2003) from US and the empirical evidence confirms the predictions of the model. First, I find that the estimated coefficients are consistent with a hump shaped evolution of interest rates through the course of a relationship. However, if I restrict the sample to those firms with higher levels of leverage (higher risk) the estimated coefficients indicate a decreasing evolution of interest rates. Similarly, if I restrict the sample to those firms with medium and low levels of leverage, the estimated coefficients again support a hump shape evolution. Furthermore, the results confirm that such hump shaped evolution is more likely to occur with asymmetry of information and that on the contrary, interest rates decrease with the length of relationship when there is symmetry of information.

The current theoretical literature about financial constraints and young/small firms financing problems does not consider how the common learning process about firms' quality, for both firms and banks, affects the cost of credit and the optimal investment decisions in the presence of financial frictions (see for instance Berger and Udell (1998) for a review of the literature about small business finance).

The first theoretical models of financial markets imperfections derived financial constraints such as credit rationing or high costs of external funds as a result of asymmetric information problems, costly state verification or contract imperfect enforceability, but they do not have precise implications about the intensity of the problem for young or small firms (any firm may suffer from these problems)[16].

Based on the importance of bank lending for small and young firms, a second and large body of the literature has focused on the evolution of interest rates or the access to credit over the time or the duration of a relationship between firms and their creditors. This literature has specialized in modeling the characteristics of loan contracts and how these characteristics may vary with the length, the scope, or the number of relationships, and with the credit markets conditions, such as the competitiveness of the markets or the size of the banks (for instance, Boot and Thakor (1994), Greenbaum *et al.* (1989), Sharpe (1990), and Wilson (1993), Von Thadden (1995), Rajan (1992), Petersen and Rajan (1995)).

Most of these models assume that the bank needs to learn or is uncertain about the firms' quality, but they do not take into account the effect of firms' learning on their financial decisions and on the observed loan characteristics. None of these theories take into account the possible effects of the asymmetry of information on the demand for capital and external funds.

This theoretical brand of the literature also generates conflicting predictions about the evolution of interest rates over the duration of a relationship. On the one hand, Boot and

[16] The most well-known models of capital market imperfections are Jaffee and Russell (1976), Stiglitz and Weiss (1981), Myers and Majluf (1984), Williamson (1987), Diamond (1989), Gale and Hellwig (1985), Townsend (1979). See Hubbard (1998) or Stein (2003) for a survey of models of investment under capital market imperfections.

Thakor (1994) show that loan rates decline as a relationship matures. This is based on a repeated moral hazard game without learning. On the other hand, Greenbaum *et al.* (1989), Sharpe (1990), and Wilson (1993) predict that loan rates increase with the duration of a bank–firm relationship. They argue that the bank's improved knowledge locks the borrower into the relationship, enabling the bank to charge above-cost interest rates as the relationship continues.

In two seminal papers that have generated a lot of posterior related research, Petersen and Rajan (1994, 1995) investigated the effect of relationship lending on the availability of credit and on the cost of credit. For their empirical strategy, they used a previous version of the dataset used here, the National Survey of Small Business Finance 1988. In their paper, they were not able to find a significant effect of the length of relationship on the interest rate, but they found a different marginal effect of the age of the firm on the interest rates depending on the level of competition in the credit market the firm is located.

Their explanation, which was supported by a theoretical model, was that the level of competition on the credit markets may affect the incentives of banks to lend to the firms. In concentrated markets a bank has an incentive to offer cheap credit at the beginning of a relationship given that because of the lack of competition the firm will be lock-in into a relationship and could be forced to pay higher interest rates after a relationship has been established.

The implications and the results presented here may look similar to those of models that predict that interest rates may increase with the duration of a relationship due to a lock in effect or an increase in monitoring costs. For instance, Kim, Kristiansen and Vale (2008) find that Banks' interest rate markups are predicted to follow a life-cycle pattern over the borrowing firms' age due to endogenous bank monitoring by competing banks. In their model, interest rates first increase with the age of the firm and decrease afterwards. The results from this last paper are similar to mine, in the sense that they may imply a similar evolution of interest rates.

However, in their theoretical model all firms are assumed perfectly informed and they do not decide about their optimal level of capital or the optimal cost of external funds. Is the competition among banks which generates that the markup interest rate is initially negative, increases afterwards, and finally decreases. In the model presented here both firms and banks estimate the riskiness of investment and is the fact that financially constrained firms choose riskier investment projects as banks learn which may generate an increase in interest rates with the length of relationship.

There are two different empirical results that support the theoretical prediction of this chapter over those based on the lock-in effect or hold-up problem. First of all, models based on the lock-in problem do not predict a different evolution of interest rates depending on whether the length of relationship is equal or smaller than the age of the firm, meanwhile the model presented here predicts that interest rates decrease with longer relationships when the length is equal to the age of the firm and a hump-shape evolution when they are different; something that is confirmed with the data. Secondly, in a lock-in problem model the evolution of interest rates should be the same independently of the level of risk of the firm (more or less leverage or internal funds), meanwhile the empirical evidence presented herein shows that the effect is indeed different.

The results of the chapter are also important because they could help clarifying why the empirical results of the different studies that have studied the relationship between interest rates on commercial loans and the length of a relationship using similar US data is mixed and inconclusive. On the one hand[17], Berger and Udell (1995) is the only paper that finds a negative and significant effect of the length of a relationship on the interest rates paid using data from the NSSBF. Their result is, however, restricted to lines of credit and floating interest rates (the effect was on the points over the respective index). They also find a significant negative effect of the length of relationship on the probability of pledging collateral.

On the other hand, several studies that use data from other countries show conflicting evidence with respect to Berger and Udell (1995). Degryse and van Cayseele (2000) find a positive effect of length on interest rates using a large survey of Belgian firms. Blackwell and Winters (1997) do not find a significant effect of the length on interest rates. They, however, find a negative effect of the length on monitoring frequency. Elsas Krahnen (1997) and Harhoff Körting (1998) neither find a significant effect of length on cost of credit. Finally, Angelini, Di Salvo and Ferri (1998) find a positive effect of length on interest rates using italian survey data.

More recent versions of the National Survey of Small Business Finance have included a much wider set of information about the firm's owners, the loans' characteristics and the relationship between firms and the financial institutions they have business with. However, to the best of my knowledge, there are no papers that have been able to find a significant effect of the length on the interest rates using the more recent versions of the NSSBF. For instance, several papers that have explored other related topics such as race and gender discrimination in the small credit market, or the distance from a bank, have not reported a significant relationship between the length of the relationship and the interest rate paid on the more recent loan using different versions of the NSSBF. (For instance, Blanchard, Levine and Zimmerman (1998), Cavalluzzo and Wolken, (2002), Petersen and Rajan (2002)). Some of these papers do not even report a significant effect of firms' age on interest rates.

Therefore, the results presented here may help to understand such conflicting evidence. More importantly, it also adds the novel result that the final effect of the length of relationship depends on the degree of asymmetric information and on the degree of a firm's riskiness. Therefore, the main added value of the empirical analysis is to show that the effect of the length of relationship on loan characteristics can only be properly understood by considering the role of internal funds and the perception of risk on the demand for debt, and the role of the length relative to the age of the firm.

This chapter is organized as follows: In section 2 I explain the relationship between relationship lending and interest rates, and I perform some theoretical simulations to

[17] According to two different surveys of the literature: Ongena and Smith (1998) and Elyasiani and Goldberg (2004).

exemplify such relationship. In section 3 I present some testable implications of the model, the empirical strategy and the estimation results. Section 4 concludes.

5.2 Effect of Relationship Lending on Interest Rates

According to section 4.2.5 which establishes how the probability of default and the required interest rate on a risky loan are defined by the creditor, the interest rate R_1 is given by:

$$R_1 = \left[\frac{RB_1}{(1-\Phi_B)B_1} - \frac{\Phi_B q k_1 (1-\tau)}{(1-\Phi_B)B_1} - \frac{(\tilde{\Theta}_{1B} \mid D) k_1^{\beta}}{(1-\Phi_B)B_1} \right]$$

The probability of default Φ_B is determined by the default point $\overline{\Theta}_1$ which is defined in equation (15). $(\tilde{\Theta}_{1B} \mid D)$ represents the partial expectation of future revenues below the default point, according to bank's information, which also depends on equation (15). B_1 represents the amount of debt needed to finance capital level k_1, which is determined according to the optimal decision of the firm.

As explained in section 4.2.6, the firm maximizes its expected value which is defined by equation (18) subject to the creditor's rule for the optimal interest rate, which is determined by equation (16). The first order condition for capital when a firm faces bankruptcy costs and asymmetry of information is given by equation (20), meanwhile the first-best level of capital is determined by equation (10).

Based on the theoretical model presented in Chapter 4, this chapter shows that there are two separated effects of a longer relationship between a bank and a firm on interest rates.

The first one is a reduction in the uncertainty about the expected firm's profitability. This effect is the same for the firm and the bank and it comes from the information that is accumulated through time since the relationship starts. This I formally proved in the following section in Proposition 8 and Proposition 9.

The second one is a reduction in the asymmetry of information and it comes from the convergence in the relative amount of information that the bank and the firm have, compared to how much information they have at the beginning of the relationship. Thus, the longer a relationship is, the higher the reduction in the possible disagreements about their estimated default probabilities and the expected values of revenues in the case the firm defaults (or repays). As N increases, so does T, and therefore, the functions that depend on the amount of information should converge to the same values for bank and firm, as is the case of Φ_B and Φ_F, and $(\tilde{\Theta}_{1B} \mid D)$ and $(\tilde{\Theta}_{1F} \mid D)$.

5.2.1 Effect of the Reduction in Uncertainty (Less Variance)

The first effect, the reduction in uncertainty is shown as a reduction in the variance of prediction of the bank (and of the firm if there is symmetry of information). The reduction in the variance of prediction always generates a reduction in the interest rate of risky loans via a reduction in the estimated probability of default Φ_B. This is formalized in the following propositions:

Proposition 8.

As the length of the relationship between a bank and a firm increases (as $N \to \infty$), the estimated probability of default Φ_B decreases, driven by the decrease in the variance of prediction σ_B.

Proof *Appendix A. 3*

Proposition 9.

The interest rate R_1 decreases with the length of relationship (and with the age of the firm) for any constant level of k_1, because R_1 is a decreasing function of σ_B, keeping k_1 constant.

Proof *Appendix A. 3*

These two propositions imply that in the symmetric case ($T = N$) and no bankruptcy costs, the model predicts that R_1 should always decrease as $N \to \infty$, given that when $\tau = 0$ the optimal value of capital is constant and equal to the First-best level.

Proposition 9 only guarantees that the interest rate decrease with less uncertainty when $\tau = 0$, but it does not guarantee that the interest rate decreases in the case of symmetry but positive bankruptcy costs $\tau > 0$, since the presence of the bankruptcy cost affects the optimal value of capital. However, it is clear that the total effect of the bankruptcy cost, which is given by $\tau\Phi_B q k_1$, is higher the higher the probability of default estimated by the bank Φ_B is, and therefore, in both asymmetric and symmetric cases the bankruptcy costs decrease with bank´s learning, because in both cases the probability Φ_B decreases as $N \to \infty$.

5.2.2 Effect of the Reduction in Asymmetry (Less Disagreement)

The reduction in the asymmetry of information and the possible disagreements between the firm and bank's estimations has an effect on the loan's interest rate via a change in the firm's demand for capital and credit.

In the previous chapter it was shown that any disagreement between bank and firm is translated into a firm's lower demand for capital (Proposition 5), i.e. in the asymmetric case investment demand is constrained. And the Corollary 1 for Proposition 1 and Proposition 2 establishes that the loan's interest rate depends on the amount of capital borrowed and on the firm's leverage, which implies that a lower demand of external funds translates into a lower interest rate. Thus, this implies that the initial effect of asymmetry is a reduction in the interest rate paid by firms at the beginning of the relationship.

However, as it is shown in the previous chapter, the effect on the demand for capital caused by a disagreement differs for firms with different amounts of internal funds W_0 and thus, with different levels of risk, as it is shown in Proposition 6 in section 4.31. Firms with intermediate-low levels of internal funds are the ones that invest at the lowest level when facing asymmetry, meanwhile firms with very low levels of internal funds invest closer to their first-best levels of capital. Thus, in Chapter 4 it is shown that the effect of asymmetry on the demand for credit actually translates into a u-shaped relationship between capital and internal funds. This implies that the demand for credit is reduced more strongly for firms with intermediate levels of risk; meanwhile the reduction in the demand for credit is much smaller for firms with a high level of risk with respect to their optimal first-best levels.

Hence, the initial effect of asymmetry (disagreement) is much stronger for firms with a low or intermediate level of risk than for firms with very high levels of risk (high leverage). What this means is that as low-risk or intermediate-risk firms initially demand much less credit, they also initially pay a much lower interest rate than high-risk firms, with respect to the initial interest rate they would pay if they demanded their first-best level of capital.

This implies that the second effect of a longer relationship on interest rates has a different impact for firms with different levels of risk. The difference comes from how much the demand for capital is affected by the initial asymmetry of information.

On the one hand, those firms for which the reduction in the demand for capital is more pronounced at the beginning of the relationship, there is a large and positive impact of a longer relationship on interest rates, because as asymmetry of information disappears they strongly increase their demand for capital and their leverage.

On the other hand, those firms for which the initial reduction in the demand for capital is less acute, the further increase in their demand for capital as the initial asymmetry vanishes is positive but small.

5.2.3 Total Effect of Relationship Lending

The total effect of relationship lending is the sum of the two previously described effects that could have different signs (the reduction in uncertainty and the reduction in disagreement). Therefore, the evolution of interest rates through the course of a relationship depends on which effect dominates.

It is likely that for those firms for which the effect on the demand for capital is small (the high risk firms) we should observe that interest rates decrease or do not change as the length of relationship increases (as $N \to \infty$).

On the contrary, it is likely that for those firms for which the reduction in the demand for capital is high enough (low and intermediate risk firms) we may observe that interest rates are initially very low (given the highly constrained demand of capital) and that they increase as asymmetry disappears.

As a way to illustrate and evaluate the implications of the two different effect of a longer relationship on interest rates I perform two different exercises using the FOC for capital to simulate the behavior of a firm and the evolution of interest rates along the course of a relationship.

In the first exercise I simulate the optimal evolution of the demand for capital and of interest rates for different firms with different amounts of internal funds, and for two different cases in which at the beginning of the relationship (*N=0*) the firms have the same information (*T=0*) or more information (*T*>0) than the bank.

For determining the different levels of internal funds for the different firms in this first exercise I use some statistic characteristics coming from the sample distribution of firms in the National Survey of Small Business Finances (NSSBF 2003). More specifically, I use the ten percentiles values of the ratio of Financial Wealth to Assets in the NSSBF sample to calculate ten different values of internal funds. The variable financial wealth is defined as the sum of the variables cash in hand, accounts receivables and inventory minus the variable accounts payable.

For this initial example, I calibrate the parameters of the model so that the first-best level of capital of the theoretical model is equal to the median value of the variable Assets in the NSSBF sample, which is equal to US$ 173.000. The parameters used are the following: the risk-free real interest rate *R* is set equal to 1.02, which is the average real return on a one-year U.S. T-bill between 1986 and 2005. The returns-to-scale parameter is set equal to 0.97. The depreciation rate is set to 10% and the percentage of tangible assets is set in 60%, therefore, the parameter *q* is equal to 50%.

The risk-free interest rate, the returns-to-scale parameter and the depreciation rate come from Caggese (2007). The percentage of tangible assets comes from Shaikh (2004) who suggests that 50-90% of a company's value is derived from its intangible assets and from Blair and Wallman (2001) who estimates that the percentage of the market value of a corporation due to tangible assets has varied from 80% to 20% from 1974 to 1998 (50% on average).

Since the first-best level is given by $k_1^* = \left[\frac{R-q}{\beta\tilde{\Theta}_{1F}}\right]^{1/(\beta-1)} = 173.000$, the value of $\tilde{\Theta}_{1F}$ must be equal to 0.7882, which implies that θ, the mean of the normal distribution, must be equal to -0.238.

I then use the distribution of the ratio of F.Wealth/Assets to calculate 10 different values of internal funds by multiplying the defined first-best value of assets by the ratio.

Table 28 Percentiles of Internal Funds according to NSBF 2003

WEALTH/ASSETS	MEDIAN ASSETS K1	F.WEALTH W0
0.9792	173000	169409
0.8279	173000	143223
0.6843	173000	118391
0.5536	173000	95776
0.4286	173000	74143
0.3073	173000	53164
0.1838	173000	31803
0.0805	173000	13931
0.0000	173000	0
-0.016	173000	-2730

I afterwards simulate the evolution of the optimal value of capital and of the interest rate for each of the ten different values of internal funds. For each value of internal funds I simulate the evolution in both the symmetric and the asymmetric case. For the asymmetric case I set the value of $N = 0$ and the value of $T = 3$, since 3 is the median age of those firms which have a relationship with a bank of less than one year in the NSSBF sample. Finally, I set the value of τ as 10%.

In Figure 21 we can observe the simulated evolution of the interest rate in the symmetric case for the different levels of internal funds. It is clear that independently of the level of risk, the interest rate always decreases with a longer length of relationship (more information and less uncertainty in the prediction)[18].

On the contrary, in Figure 22 we can observe that the interest rate increases with the length of relationship for the firms with the five highest levels of financial wealth (from 95.776 to 13.931) in the asymmetric cases ($T = 3$, $N = 0$). However, interest rate R_1 still decreases for the two firms with the lowest levels of financial wealth (worst risks).

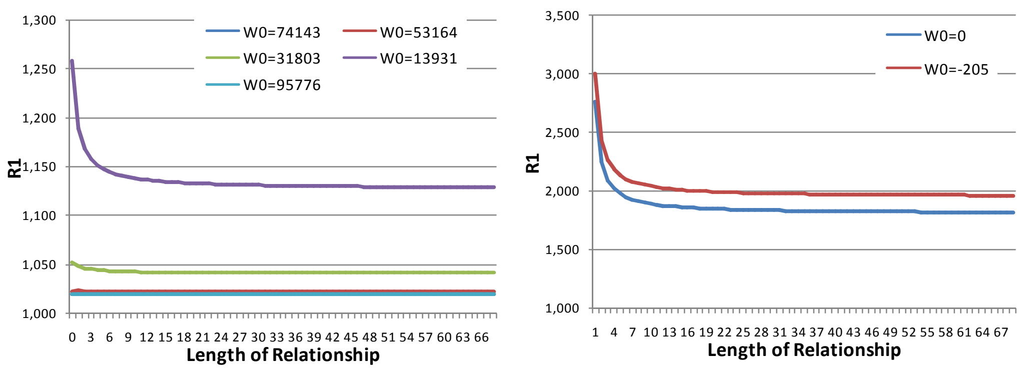


Figure 21. Evolution of interest rate R_1 with the length of relationship, for different values of internal funds and risk in the Symmetric case

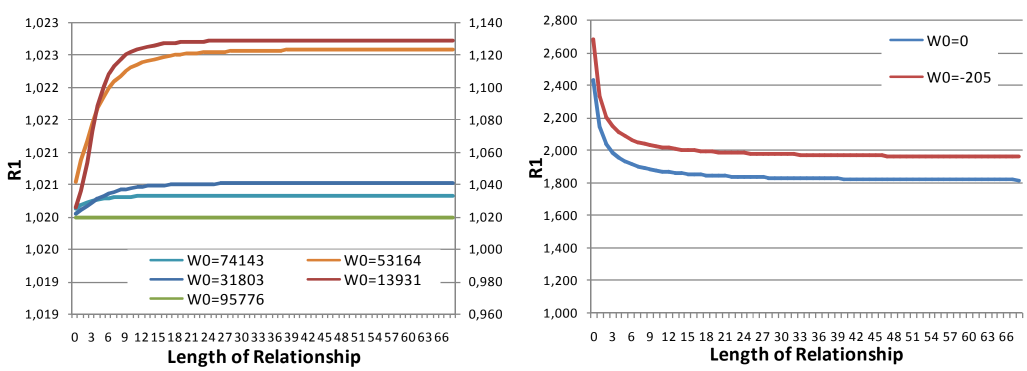


Figure 22. Evolution of interest rate R_1 with the length of relationship, for different values of internal funds and risk in the Asymmetric case

[18] In Figure 21 andFigure 22 I do not show the curves for the firms with the higher levels of internal funds, since their levels of investment are unconstrained and they pay the risf-free rate.

In the second simulation exercise I generate a random sample of 2000 firms based on the distribution of the variables Financial Wealth/Assets, Age and Length/Age in the NSBBF 2003. Since the most important exogenous variables in the model are internal funds, the age of the firm, the length of the relationship and the relationship between the age and the length (symmetry vs. asymmetry), I use the statistic distribution characteristics of these three variables in the sample of the NSSBF 2003 to generate an random sample of firms that replicates the characteristics observed in those three variables.

For the internal funds I use the same percentile values presented before and I generate 200 random observations for each range composed for the ten percentiles values calculated in the first exercise and showed in Table 28. Therefore I generate 200 random values from -205 to 0, 200 random values from 0 to 13931, etc. according to a uniform distribution inside each range. Therefore I finally have a random sample of 2000 values of internal funds that follows the same statistical distribution of the variable Financial Wealth/Assets in the NSSBF 2003.

I also compute the 4 quartiles of the variable AGE in the NSSBF and generate 400 random values for each of the 5 ranges composed by the 4 quartiles and the minimum and maximum values, as shown in Table 29.

Afterwards, for each range of the variable Age I compute the quartiles of the variable Length/Age according to the sample of NSSBF 2003. Therefore, I compute the distribution of the variable length conditional on the values of the variable age, as shown in Table 29.

I use the conditional quartiles of the variable Length/Age to compute 80 random values for each of the ranges of the variable Length. Therefore, I have 400 random values of the variable Length for each range of the variable Age.

Finally, each observation in the simulated sample is composed of 3 randomly generated numbers. The first two numbers are a duple of Age and Length and the third number is one of the random values of internal funds.

Table 29 Ranges for random sample generation based on the quantiles of AGE and LENGTH/AGE

		VALUES FOR RANGES OF LENGTHMAGE					
RANGES OF AGE	*0 to 8*	0	0.65	0.85	1	1	1
	9 to 14	0	0.5	0.75	0.91	1	1
	15 to 20	0	0.44	0.68	0.85	1	1
	21 to 28	0	0.35	0.57	0.78	0.95	1
	29 to 60	0	0.3	0.47	0.63	0.86	1

I use the FOC for the demand for capital established in equation (20) and the random generated values of internal funds, age and length to calculate the values of capital, debt and interest rates paid by each firm in the simulated sample. The interest rate is calculated using equation (16).

I finally run a simple regression with the interest rate R_1 as dependent variable, the log of age, the log of length, the value of debt B_1 and the value of optimal capital k_1 as independent variables. The results of this regression are shown in Table 30 in the column named "Whole Sample". In the column "High Risk" I show the results of the regression restricting the sample to the firms with the lowest levels of financial wealth. In the column "Low Risk" I show the results of the regression restricting the sample to the firms with the highest levels of internal funds.

In the column "Symmetry" of Table 30 I show the results of restricting the sample to firms for which their length is equal to their age. In the column "Asymmetry" of Table 30 I show the results of restricting the sample to firms for which their ratio of Length/Age is less than 0.9 (the ones with a higher asymmetry).

Finally, I further split the sub-samples "Symmetry" and "Asymmetry" into smaller subsamples of "Low Risk" and "High Risk" firms and run the same regression specification using these four subsamples. The results of these regressions are shown in Table 31.

It can be seen in the first column of Table 30 (whole sample) that both the logs of age and the length of relationship are statistically significant, although they show opposite signs. The effect of the log of age is negative and the effect of the log of length is positive. An important result is that both of them are significant despite that I am controlling for the level of capital and the level of debt.

The predicted evolution of interest rates given by the combined effects of the age and the length of relationship in the first regression in Table 30 is shown in Figure 23 (Whole sample). I use the estimated coefficients of the different regressions to simulate the evolution of interest rates of a firm that has initially an equal age and length ($T = N = 0$) and of a firm that has initially an age larger than zero at the beginning of the relationship ($T = 5$, $N = 0$)

In Figure 23 we can observe the simulation results. They indicate that on average we should observe that interest rates strictly decrease after the start of a relationship if there is symmetry of information, and that interest rates should follow a hump shaped evolution in the case of asymmetry of information.

On the one hand, the strictly decreasing evolution in the symmetric case comes from the fact that the estimated coefficient for the log of age is larger than the one for the log of length and therefore, if the length and the age are always the same, the sum of both coefficients is always negative.

On the other hand, when the length is shorter than the age (asymmetry), the final evolution of interest is more involved. Since the marginal effect of a variable in logs is higher the smaller is the variable, we have that the marginal effect of 1 additional year of relationship when the length is 0 could be much higher than the effect of 1 additional year of age is the age is, for instance, 5 years old. This generates an initial increase in interest rates just after a relationship starts. This is so despite the coefficient on LN(LENGTH) (1,3343) is smaller than the coefficient of LN(AGE) (-1,7312) in absolute terms.

However, if we restrict the sample to the case of firms with the lowest levels of internal funds (High Risk) in Table 30 we can observe that the sign of both the log of age and the length switch. Furthermore, if we look at the regression results that use the subsamples "High Risk - Asymmetric" and "High Risk - Symmetric" in Table 31 and the implied evolution shown in Figure 25 (High Risk subsamples), the estimated coefficients predict that the evolution of interest rates is always negative independently of whether there is symmetry or asymmetry of information, in the "High Risk" case.

Finally, if we restrict the sample to the firms with the highest levels of internal funds (Low Risk), in Table 30 we can observe that the sign of both variables change again and are similar to the ones observed for the whole sample, although the difference between them is smaller. Additionally, in Table 31 we can observe that in the "Low Risk" subsamples the sign of the coefficient of the variable LN(LENGTH) is positive and significant in the "Asymmetric" subsample and negative and significant in the "Symmetric" one. Therefore, the results differ from the previous case, as shown in Figure 24 (Low Risk). The estimated coefficients predict that the evolution of interest rates is positive in the "Low Risk – Asymmetric" subsample and the evolution is negative in the "Low Risk – Symmetric" one.

Table 30 Simulation Regression Results

	Whole Sample	High Risk	Low Risk	Symmetric	Asymmetric
R2 =	0.3297	0.8950	0.5033	0.3871	0.3157
Adj. R2 =	0.3283	0.8942	0.5020	0.3810	0.3134
F-stat	235.371	1052.895	358.337	63.155	137.356
Obs. =	2000	665	1271	344	1419
	Whole Sample	**High Risk**	**Low Risk**	**Symmetric**	**Asymmetric**
LNLENGTH	1.3344***	-9.8951***	0.9737***	-1.2352*	1.2022**
	(0.000)	(0.000)	(0.000)	(0.085)	(0.045)
LNAGE	-1.7312***	4.3864***	-0.8135***		-1.0964*
	(0.000)	(0.000)	(0.000)		(0.095)
K1	-0.0186***	-0.2646***	-0.0054***	-0.0282***	-0.0169***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
B1	0.01934***	0.3342***	0.0001	0.0099**	0.0224***
	(0.000)	(0.000)	(0.645)	(0.047)	(0.000)
Intercept	24.4866***	16.5450***	13.6401***	42.4157***	19.5604***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

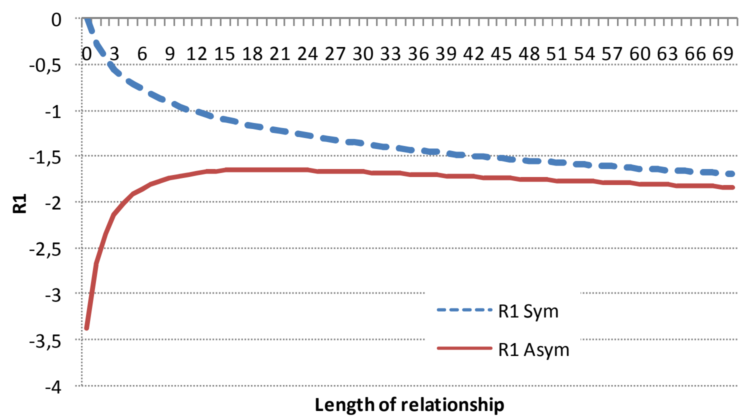


Figure 23 Predicted evolution of interest rate R_1 in the whole sample of simulated firms. Symmetric vs Asymmetric cases

Table 31 Simulation Regression Results. Different sub-samples.

	Low Risk	Low Risk	High Risk	High Risk
R2 =	0.660	0.786	0.858	0.913
Adj. R2 =	0.659	0.783	0.857	0.910
F-stat	409.5	214.87	770.12	421.13
Obs. =	884	187	535	130
	Asymmetric	**Symmetric**	**Asymmetric**	**Symmetric**
LNLENGTH	0.269***	-0.112**	-17.846***	-8.636***
	(0.000)	(0.014)	(0.000)	(0.000)
LNAGE	-0.179***		9.919***	
	(0.000)		(0.000)	
K1	-0.0023***	-0.004***	-0.1137***	0.080**
	(0.000)	(0.000)	(0.000)	(0.001)
B1	-0.0003**	-0.0007***	0.2336***	0.3358***
	(0.002)	(0.000)	(0.000)	(0.000)
Intercept	9.293***	12.308***	-23.760***	-55.694***
	(0.000)	(0.000)	(0.000)	(0.000)

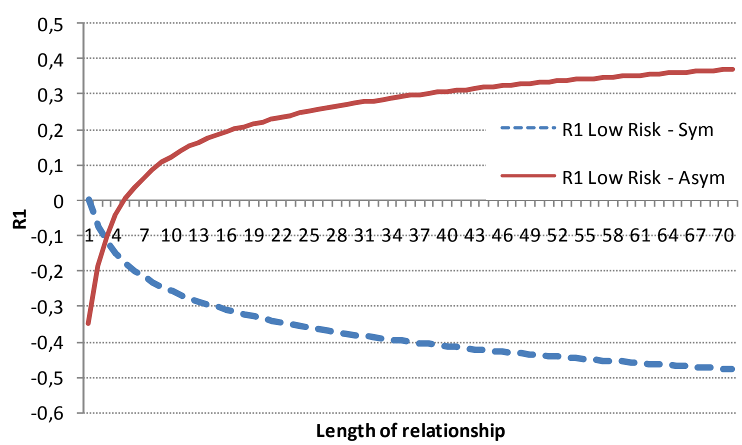


Figure 24 Predicted evolution of interest rate R_1 in the sample of Low Risk simulated firms. Symmetric vs Asymmetric cases

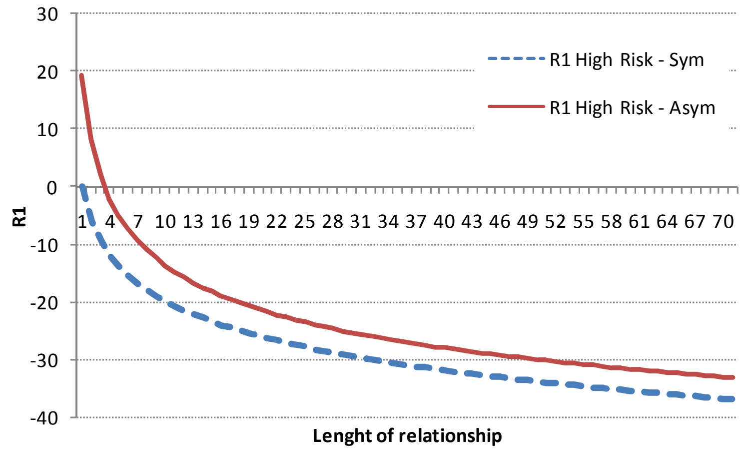


Figure 25 Predicted evolution of interest rate R_1 in the sample of High Risk simulated firms. Symmetric vs Asymmetric cases

The type of evolution shown in Table 30, Figure 23, Figure 24 and Figure 25 is what we should expect to find as the effect of the age and the length of relationship on interest rates. According to Figure 23, we should find that, on average, with symmetry of information interest rates decrease monotonically as the length increases and that, on average, with asymmetry we should observe a hump shaped evolution. On the contrary, we should find that interest rates decrease or do not change after the start of a relationship in the riskier firms' case. Additionally, we should find that interest rates increase for firms facing asymmetry of information in the less risky case and that on the contrary, they decrease if there is symmetry of information[19].

[19] In the previous simulations, I keep the value of internal funds W_0 constant. But it is clear that in a more dynamic framework the initial value of internal funds at each point in time may change with the age of the firms when firms are financially constrained. If firms are constrained, they may accumulate internal funds with time to reduce the riskiness and the cost of capital, and therefore, W_0 may increase with the age of the firm. In such case the evolution of interest rates may look a bit different than what is shown in these graphs.

5.3 Dataset and Empirical Investigation

The dataset used is the 2003 National Survey of Small Business Finances (NSSBF), the last version of a series of surveys conducted by the Board of Governors of the Federal Reserve and the U.S. Small Business Administration. The survey collects information on the use of financial services and institutions for a US nationally representative sample of firms. The target population is all nonfinancial, nonfarm small business firms with fewer than 500 employees.

Different versions of this survey have been extensively used to analyze the financing patterns of small firms, including some of the most well-known empirical analysis of relationship lending (Petersen and Rajan (1994, 1995), Berger and Udell (1995) among others).

The sample design (stratified random sample with oversampling of certain groups) used in the NSSBF 2003 requires using special complex sample techniques. As noticed in the survey codebook "Ignoring the survey aspects of the data will generally lead to incorrect point and parameter estimates and generally underestimated estimated variance and standard errors, sometimes substantially so". The statistical software that was used for the analysis was STATA. This statistical software requires that there should be more than one observation per each strata unit (the same as most of the statistical packages that work with complex survey data). To handle this kind of problem I chose to drop those strata with single observations when this problem appeared. This implied losing some observations, and may imply that point estimations may differ a little bit depending on the statistical package used or in the way of handling this problem.

The most important variables for the analysis are the length of the relationship and the age of the firm. In different studies of relationship lending, different definitions of the length of relationship have been used. For instance, Petersen and Rajan (1994,1995) used the maximum length reported by a firm for their analysis of the availability of credit, meanwhile they used the reported length of relationship with the financial institution that granted the loan for their analysis of the cost of credit.

If we closely analyze the reported length of relationship for the most recent loan applied or denied, and the maximum length reported by the firm with any financial institution, we can see that in about 80% of the cases they both correspond to the length of relationship with the same primary institution. Therefore, I chose to use the maximum length reported by the firm with any financial institution as the appropriate measure of how much information a creditor can observe from a firm. This variable is called LENGTHM which correspond to the longest length of a relationship with any financial institution (in years) at the time of application. I checked the robustness of the analysis if I use the length of relationship with the institution that granted the loan instead of LENGTHM and the results remain the same.

The most important problem observed in the original variables LENGTHM and AGE in the dataset is that there are an important number of firms that report a primary relationship at the time of application longer than the age of the firm. This data issue and its possible consequences have not been analyzed by previous empirical studies of relationship lending that have used the length as an explanatory variable using the NSSBF. Given this problem, it is not possible to test exactly the kind of asymmetry proposed in the theoretical model. In the theoretical framework, the length of the relationship is by definition, smaller or equal than the age of the firm (since it was established). Or in other words, in the theoretical framework, $T>=N$, or the ratio of the length of relationship to the age of the firm should be less or equal than one.

One of the main reasons of this problem is that given the definition of a firm's age in all the versions of the NSSBF, it is not possible to know the age of a firm since it was established, but only since the current owner is the proprietorship of the firm (variable AGE). Therefore, the most obvious way to deal with this problem is to restrict the analysis only to the case of firms that were established by the current owners, which eliminates the problem of relationships longer than the age because of mergers and acquisitions.

However, around 20% of firms from the remaining observations still reported maximum lengths of relationship longer than the age of the firm. This could be a consequence that for many small firms the owner-bank relationship cannot really be distinguished from the firm-bank relationship, and many owners may report their personal relationships with their bank as a firm-bank relationship, eventhough the owner-bank relationship could have started before the start-up. Given the information in the survey and the questionnaire, it is not possible to give an exact reason of why there are lengths that are longer than the reported age.

A second possible solution is to eliminate all the observations with values of LENGTHM higher than AGE (values of the ratio of LENGTHM/AGE higher than one). I estimated the regressions with such restricted sample and all the main results presented here do not change. However, I choose not to restrict the sample in this way because we do not know the exact reason of the presence of those observations and because of a large number of observations would be lost. If I use only the sample of firms originally established by their owners and with lengths of relationship smaller or equal than the age the sample reduces to only 860 firms. Moreover, if the reason why some firms have longer lengths than its age is because they are reporting an owner-bank relationship previous to the start-up, we cannot be sure that a firm reporting a length to age ratio of 0.8 is also not reporting, for instance, the existence of a current account on the name of the owner and not in the name of the firm.

Nonetheless, I drop from the analysis those firms with clear outliers in the ratio of LENGTHM to AGE in order to bring the relation between the variables LENGTHM and AGE as close as possible to the theoretical definition. More specifically I drop from the analysis those firms with values of the ratio of LENGTHM/AGE higher than 5.[20] This is justified by the fact that, for instance, there are several firms in the sample that report maximum lengths of more than 20 years and firm ages of 0 or 1 years, which is very difficult to interpret. I check the robustness of the results to restricting the sample to firms with ratios of LENGTHM to AGE smaller than 5, 4, 3 and 1, and in all the cases the results do not change.

The full sample size contains 4240 firms. I do not use observations from firms for which the type of institution corresponding to the last application was a Government Agency, or Family or Other Individuals. From these 4240 firms, the sample of firms that reported that they

[20] which is the 99 percentile in this variable.

asked for a loan recently and that report the interest paid for it contains 1668 firms. The final sample contains about 1060 after selecting only those firms that were originally established by their current owners and after dropping outliers in the LENGTHM/AGE ratio.

5.3.1 Analysis of Cost of Credit

The dependent variable used in this part of the analysis is the interest rate quoted on the firm's most recent loan. According to the theoretical framework, we should observe a negative relation between the age of the firm and the interest rate when there is symmetry of information, given that when $T = N$ the higher the age of the firm the more information that both bank and firm have, and thus, the lower the variance of prediction and the lower the estimated probability of default (and the estimated bankruptcy cost). However, when there is asymmetry of information, i.e. when $T > N$, there could be a non-monotonic evolution of interest rates after a relationship starts. Additionally, the evolution in the case of asymmetry depends on the riskiness of the firm, and more specifically, on the level of internal funds or on leverage.

The reason for this is that the total effect of relationship lending on interest rates is the sum of two different effects that could have different signs: the reduction in uncertainty and the reduction in the disagreements between the firm and the bank. The evolution of interest rates through the course of a relationship depends on which effect dominates. According to the model simulations, for those firms for which the effect on the demand for capital is small (the high risk firms) we should observe that interest rates decrease or do not change as the length of relationship increases (as $N \to \infty$). On the contrary, for those firms for which the reduction in the demand for capital is high enough (low and intermediate risk firms) we should observe that interest rates are initially very low (given the highly constrained demand of capital) and that they increase as asymmetry disappears.

Other possible explanations for a hump shape evolution of interest rates after the start of a relationship are related to lock-in effects. Greenbaum *et al.* (1989), Sharpe (1990), and Wilson (1993) predict that loan rates increase with the duration of a bank–firm relationship. They argue that the bank's improved knowledge locks the borrower into the relationship, enabling the bank to charge above-cost interest rates as the relationship continues.

Petersen and Rajan (1994, 1995) investigated the effect of relationship lending on the availability of credit and on the cost of credit. They were not able to find a significant effect of the length of relationship on the interest rate paid by the firms in the NSSBF 1998 survey, but they found a different marginal effect of the age of the firm on the interest rates depending on the level of competition in the credit market the firm is located. Their explanation, which was supported by a theoretical model, was that the level of competition on the credit markets may affect the incentives of banks to lend to the firms. In concentrated markets a bank has an incentive to offer cheap credit at the beginning of a relationship given that because of the lack of competition the firm will be lock-in into a relationship and could be forced to pay higher interest rates after a relationship has been established.

We could also find a positive relationship between interest rates and the length of a relationship due to an increase in monitoring costs arising from a lock-in effect as it is proposed by Kim, Kristiansen and Vale (2008). However, models based on the lock-in problem do not predict a different evolution of interest rates depending on whether the length of relationship is equal or smaller than the age of the firm, meanwhile the model presented here predicts that interest rates decrease with longer relationships when the length is equal to the age and a hump shape evolution when they are different. More importantly, in a lock-in problem model the evolution of interest rates should be the same independently of the level of risk of the firm (more or less leverage or internal funds), meanwhile the empirical evidence shows that the effect is indeed different.

Summarizing, in this section I test if the suggested prediction of the model about interest rates, age of the firm, and the length of relationship can be found in the data, focusing on the different evolution in firms with different levels of risk and asymmetry, which differentiate the model results from other arguments. Thus, I test if the type of evolutions depicted in Figure 21 to Figure 25 can be observed in the real data.

Some preliminary evidence in favor of those predictions can be found in some simple summary statistics from the NSSBF 2003. In Table 32 I have calculated the average interest rate in the NSSBF for the 4 different ranges constructed according to the quartiles of the variables AGE and LENGTHM and the average for the combined conditional ranges. Younger and shorter relationships are classified in the range number 1 and older firms and longer relationships in the range 4. We first can observe that the average interest rate seems to decrease with both AGE and LENGTHM, although the age effect seems stronger.

However, we can see that the lowest average interest rate is observed in the group formed by the oldest age (range 4) but the shorter length (range 1), which therefore corresponds to the group with the higher asymmetry of information. On the contrary, the highest average interest is observed in the group of younger firms (range 1) but with the longest length in their range (range 2).

Table 32

Quartiles AGE	Quartiles LENGTHM				
	1	2	3	4	Total
1	6.812	7.310			6.798
2	6.094	6.264	6.258		6.274
3	5.367	5.666	5.760	5.345	5.597
4	4.203	5.081	5.071	5.574	5.246
Total	6.269	6.077	5.722	5.626	5.945

In Table 33 we can also observe the average interest rate for the 4 ranges based on the 3 quartile of the ratio of LENGTHM/AGE. In this case is clear that the average interest rate is higher the larger the ratio is, that is, the more symmetry of information.

Table 33

	Quartiles LENGTHM/AGE
1	5.379
2	5.689
3	6.052
4	6.590
Total	5.945

Hypotheses to test:

- *Interest rates should decrease monotonically after the start of a relationshp in the case of symmetry of information (as both the length and the age of the firm increase), but there may be a hump shaped evolution in the case of asymmetry.*

- *The marginal effect of the length of relationship should be positive when T>N for firms that show a low leverage ratio (low risk) and negative when T=N.*

- *The marginal effect of the length of relationship should be negative when T>N for firms that show a high leverage ratio (high risk) and negative when T=N as well.*

To test these predictions I estimate the following basic specification:

Interest Rate = β_0 + β_1 *LN(LENGTHM)* + β_2*LN(AGE)* + β_3 *Relationship Characteristics* + β_4 *Firm Characteristics* + β_5 *Owner Characteristics* + β_6 *Loan Characteristics+ u*

In the first and second columns of Table 34 I show the results of this regression specification in two different samples. In the first column I show the results of the regression using the original sample without any restriction ("Whole sample" column). In the second column I show the regression results after restricting the sample to firms originally established by the owners and after droping firms with clear outliers in the ratio of LENGTHM to AGE. The evolution of interest rates implied by the estimated coefficient in the restricted sample is depicted in Figure 26 ("Originally owned sample").

Next, I estimate the same regression equation but restricting the sample to firms for which the leverage ratio is lower than 0.7 (lower risk), which corresponds to 75th-percentile value. This is the sample of "Low-risk" firms. The results of this test can be seen in column 3 of Table 34, and the implied evolution of interest rates is depicted in Figure 27.

I then restrict the sample to firms for which the leverage ratio is higher than 0.7 (higher risk). This is the sample of "High-risk" firms. The results of this test can be seen in column 4 of Table 34, and the implied evolution of interest rates is depicted in Figure 28.

I also estimate the same regression but restricting the sample to firms for which AGE>LENGTHM (*T* > *N*) (Asymmetric). The results can be seen in column 5 of Table 34.

The results of the same regression but restricting the sample to firms for which AGE=LENGTHM ($T = N$) (Symmetric) can be seen in column 6 of Table 34.

Finally I further separate the Asymmetric and Symmetric subsamples into four subsamples of High Risk and Low Risk firms. The results of this last estimation exercise can be seen in Table 35.

Although not reported, I test whether the competitive conditions of the credit market may have an effect on the results. Therefore, I compare the effect of LN(LENGTHM) in the concentrated market with its effect in competitive and intermediate markets through the inclusion of interaction variables. I also test whether the effect of LN(AGE) in concentrated markets is different than its effect in competitive or intermediate markets. None of the interaction variables included were significant, showing that the results presented here are not related to the lock-in effects considered by Petersen and Rajan (1994,1995).

In the regression analysis, I control for many different observable firm characteristics and also for different aspects of a relationship that have been suggested in the literature, such as the number of financial institutions that the firm is dealing with, the owner characteristics and loan characteristics. Owner characteristics, loan characteristics and some of the relationship characteristics variables controlled here were not used for Petersen and Rajan (1994,1995) simply because they were not available in the first version of NSSBF.

The most important control variables are defined and described in Appendix A. 6.

In the set of relationship characteristics I control for the number of institutions the firm has loans (NINSTLOANS), the number of institutions the firm has a checking or savings account (CHESAV), the distance to the financial institution (DISTANCE), whether the institution granting the loan is a bank (BANK), or a non-financial institution (NONFINANC).

In the set of firm's characteristics I control for the size of the firm (LN(ASSETS)) and for the amount of money lent in the most recent loan (LN(AMOUNT)), for the firm credit score (A0_DB_CREDRK), for the increase in sales in the last 3 years (INCRSALES) for the profitability of the firm (PROFITABL), if the firm has been audited, if the firm is managed by the owners (OWNMANAGE) and for a set of 7 industry SIC categories and 4 type of corporation categories.

In the set of owners characteristics I control for the age of the owner (LNOWNAGE), the previous experience before starting the firm (PREVEXPER), the number of owners (OWNERS), the owners education (CF_EDUC), the percentage of female ownership (CF_FEMALE), the percentage of a minority race ownership (CF_MINOR), if the owner owns a home residence (OWNHOME), and for the owner's equity different from the firm or her home (OWNOWORTH).

I also control for whether the owner or the firm have been recently delinquent (DELINQU), if the firm or the owner have gone bankrupt (BANKRPT) and if the owner or the firm are in a judgment process (JUDGMT).

In the set of loan's characteristics I control for whether the interest rate is fixed or variable (FIXED), for the type of index (PRIME01), for the prime rate at the month the loan was granted (PRIMERATE), for whether the loan has been collateralized, guaranteed or has to have a balanced account (COLLAT, GUARANTEE, BALANCE), for six different categories of collateral, for six different types of loans (new lines of credit, renewals of lines of credit, vehicle loans, capital leases, equipment loans or other types). I finally control for 9 different regional categories (dummies)

5.3.2 Regression Results

In Table 34 (Whole sample column) we can see the results of estimating the basic specification using the whole sample of firms without any restriction. If I use the original sample, the estimated coefficient of the variable LENGTHM is clearly non-significant.

On the contrary, if I restrict the sample to firms originally established by their owners and dropping outliers, both the length of the relationship and the age of the firm are significant and with opposite signs. It is very important to notice that the marginal effect of LN(LENGTHM) is positive and significant. To the best of my knowledge, the length of a relationship has never been reported as positive and significant in a regression with interest rates using any version of the NSSBF. It has been reported as positive and significant in European surveys, but not using this US data. The effect of the age of the firm is negative and significant and higher in absolute terms than the marginal effect of the maximum length of relationship.

An important implication of these results is that the same estimated coefficients could indicate different evolutions of the interest rates depending on whether there is symmetry (LENGHTM=AGE, $N=T$) or asymmetry (LENGHTM<AGE, $N<T$), as predicted by the theoretical framework. The easiest way to see the combined effect of the two variables in the evolution of interest rates is to use the estimated coefficients to simulate the evolution of interest rates of a firm that faces symmetry of information versus a firm that faces asymmetry of information at the beginning of the relationship, as it is shown graphically in Figure 26. We can see that the predicted evolution of interest rates shows a very similar pattern as the one shown in Figure 23 that is based on an artificially simulated sample according to the theoretical model.

In the symmetric case ($T = N = 0$), LN(AGE) is exactly equal to LN(LENGTHM) and as the time goes by (as both the age and length increases) interest rates should decrease because the marginal effect of age (-0,6480) is higher than the marginal effect of the length (0,4003) in absolute terms.

However, in the asymmetric case, if the length is 0 and a firm is already established ($T=5 > N=0$) we have different effects when the firm starts a relationship, according to the estimated coefficients. First of all, the interest rate may increase initially. Since the marginal effect of a variable in logarithm is higher the smaller is the variable, we have that the marginal effect of 1 additional year of relationship when the length is 0 could be much higher than the effect of 1 additional year of age is the age is, for instance, 5 years old. This generates an initial increase in interest rates just after a relationship starts. This is so despite the coefficient on LN(LENGTHM) (0,3303) is smaller than the coefficient of LN(AGE) (-0,602) in absolute terms.

It is important to notice that in some previous versions of this empirical exercise I was not taking into account the presence of outliers in the relation between the variables LENGTHM and AGE. Instead, I was using the ratio of LENGTHM/AGE as a control variable. In such version, the implied evolution of interest rates was the same. If I do not control for the ratio LENGHTM/AGE or if I do not drop from the regression the firms with extreme values in this ratio, the variable LENGTHM is not significant. To find a significant effect of this variable I should control for the two described problems together (not originally owned firms and outliers).

What this indicates is that is very likely that the reason why previous studies find such mixed evidence about the effect of the length of relationship was mainly because the relationship between the variables length and age was not well defined. If we include in the analysis firms that were not established by their current owners, the length of relationship does not reflect the true degree of relationship or information that a bank has with a firm. Furthermore, it is very difficult to interpret the effect of a very long relationship in the case of a start-up firm or a firm of only one or two years old. When I correct these two problems in the sample, the effects of the length and the age are not only clearly significant, but also as predicted by the theoretical model.

In order to check the predictions of the theoretical framework, I also perform the previous regression but splitting the sample into firms with lengths smaller than their age (LENGTHM/AGE<0.9 or "Asymmetric") and firms with lengths similar to their age (0.9<LENGTHM/AGE<1.2 or "Symmetric")[21]. The results using the "Asymmetric" sample are shown in column 3 of Table 34. When I restrict the sample to those firms for which LENGTHM<AGE the marginal effects of LN(LENGTHM) and LN(AGE) remain similar to those of the sample of "originally owned" firms, that is, the evolution of interest rates is hump shaped. The results using the "Symmetric" sample are shown in column 4 of Table 34 and we can see that in this case the coefficient of LN(LENGTHM) becomes negative and significant, as implied by the theory.

The predictions of the theoretical framework are also confirmed in the regressions in which I split the sample according to according to the firms' leverage (risk). In the columns "Low Risk" and "High Risk" of Table 34 we can see that if I split the sample into two subsamples with the firms with the highest leverage vs. the firms with intermediate or low leverage, the results change in the way predicted by the theoretical framework.

On the one hand, in the "Low Risk" sample both the coefficients of LN(LENGTHM) and LN(AGE) become larger (more positive) than in the "Originally owned" sample. The coefficients in this case imply that the evolution of interest rates is still hump shaped in the asymmetric case and negative in the symmetric case. On the other hand, in the "High Risk" regression the length is no longer significant, but the sign of the age is negative and

[21] I have also perform the "Symmetric" regression using only firms with LENGHTM=AGE and the results are the same.

significant, which implies that the evolution of interest rates is always negative in both the asymmetric and symmetric cases, as it was found in the simulation results.
The predicted evolution of interest rates according to the degree of asymmetry and the level of risk is further confirmed in the regression results shown in Table 35, in which I split the two subsamples used in the "Low Risk" and "High Risk" cases into smaller subsamples of asymmetric and symmetric cases.

If we look at the regressions using the subsamples "Low Risk – Asymmetry" and "Low Risk – Symmetry" that are shown in columns 1 and 2 of Table 35 the evolution in the "Low Risk – Asymmetric" case is hump shaped, whereas in the "Low Risk – Symmetric" case is negative (although non-significant). These predicted evolutions are shown in Figure 27 and if they are compared to the theoretical evolution depicted in Figure 24, we can see that they are very similar. In the same fashion, if we look at the regressions using the subsamples "High Risk – Asymmetry" and "High Risk – Symmetry" we can see that the estimated coefficients imply that interest rates decrease with the length of relationship in the symmetric case, and that interest rates are non-significantly correlated to the length or the age in the asymmetric case. If we compare the theoretical evolution to the empirically estimated one, we can see that, again, they look pretty similar, as we can see if we compare Figure 25 versus Figure 28.

Table 34 Complex Survey Regression Results. Split Sample Results

	Whole Sample	**Originally Owned**	**Asymmetric**	**Symmetric**	**Low risk**	**High risk**
	Number of obs = 1668	Number of obs = 1040	Number of obs = 531	Number of obs = 435	Number of obs = 626	Number of obs = 414
	F(60, 1607) = 9.25	F(60, 979) = 6.79	F(60, 470) = 5.56	F(59, 375) = 4.12	F(60, 565) = 7.45	F(60, 353) = 2.83
	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000
	R-squared= 0.2614	R-squared = 0.2750	R-squared = 0.3410	R-squared = 0.3516	R-squared = 0.3224	R-squared = 0.3525
	Whole Sample	**Original Owned**	**Asymmetric**	**Symmetric**	**Low risk**	**High risk**
LN(LENGTHM)	0.0449	0.3303**	0.4018**	-0.6634***	0.4602***	0.1688
	(0.627)	(0.031)	(0.039)	(0.005)	(0.008)	(0.574)
LN(AGE)	-0.2934**	-0.6019***	-0.6573**		-0.5529**	-0.5881*
	(0.011)	(0.001)	(0.020)		(0.021)	(0.078)
LN(AMOUNT)	-0.3057***	-0.4093***	-0.3674***	-0.5241***	-0.4549***	-0.2902**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.015)
LN(ASSETS)	-0.1018**	-0.0312	-0.0715	-0.0107	-0.0308	-0.0915
	(0.050)	(0.656)	(0.339)	(0.934)	(0.714)	(0.455)
DISTANCE	0.0008	0.0011	0.0015**	0.0000	0.0025***	-0.0007
	(0.178)	(0.152)	(0.019)	(0.994)	(0.009)	(0.345)
CHESAV	0.0663	-0.1185	0.0000	-0.1714	-0.1195	-0.1714
	(0.574)	(0.434)	(1.000)	(0.501)	(0.533)	(0.526)
NINSTLOANS	0.0469*	0.0127	-0.0430	0.1278*	-0.0389	0.0548
	(0.097)	(0.754)	(0.367)	(0.094)	(0.560)	(0.378)
ONEBANK	-0.1363	-0.4126*	-0.2777	-0.7779*	-0.2367	-0.9179**
	(0.437)	(0.087)	(0.379)	(0.052)	(0.445)	(0.045)
BANK	-0.2915	-0.2556	0.1352	-0.0475	0.3807	-0.9347*
	(0.277)	(0.468)	(0.734)	(0.929)	(0.357)	(0.078)
INTERM	-0.3643	-0.4808	-0.3543	-1.2729*	-0.4569	-0.5967
	(0.140)	(0.180)	(0.304)	(0.060)	(0.388)	(0.165)
CONCENTR	-0.1816	-0.3307	-0.4390	-0.8680	-0.4727	-0.4807
	(0.483)	(0.381)	(0.206)	(0.208)	(0.381)	(0.299)

***,**,* indicate significance at 1%, 5% and 10% respectively. P-values are in parenthesis. The complete regression results are shown in Appendix A. 7

Table 35 Complex Sample Regression Results. More Restricted Samples

	Low Risk		High Risk	
	Number of obs = 334	Number of obs = 249	Number of obs = 199	Number of obs = 186
	F(59, 273) = 5.32	F(58, 190) = 5.61	F(59, 138) = 4.98	F(58, 127) = 4.24
	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000
	R-squared = 0.4139	R-squared = 0.4157	R-squared = 0.4720	R-squared = 0.5553
	Low Risk - Asymmetric	**Low Risk - Symmetric**	**High Risk - Asymmetric**	**High Risk - Symmetric**
LNLENGTHM	0.4951**	-0.4585	0.2257	-0.9689**
	(0.025)	(0.189)	(0.586)	(0.012)
LNAGE	-0.6804**		-0.6567	
	(0.023)		(0.293)	
LN(AMOUNT)	-0.4352***	-0.4557**	-0.1617	-0.4358**
	(0.003)	(0.012)	(0.310)	(0.032)
LN(ASSETS)	-0.0611	-0.1928	-0.2377*	-0.0095
	(0.528)	(0.294)	(0.082)	(0.961)
DISTANCE	0.0019**	0.0000	0.0003	-0.0002
	(0.038)	(0.989)	(0.695)	(0.883)
CHESAV	0.1041	-0.3569	-0.2744	0.1357
	(0.675)	(0.299)	(0.448)	(0.793)
NINSTLOANS	-0.1102	0.0333	0.0332	0.2672*
	(0.163)	(0.769)	(0.714)	(0.070)
ONEBANK	-0.1868	-1.2177**	-1.2567	0.3611
	(0.605)	(0.030)	(0.138)	(0.588)
BANK	0.8456	0.3416	-0.4565	-0.6698
	(0.137)	(0.603)	(0.403)	(0.401)
INTERM	0.0716	-2.3028**	-1.1176*	-0.6874
	(0.883)	(0.048)	(0.081)	(0.368)
CONCENTR	-0.2178	-1.7556	-0.7159	-1.0351
	(0.652)	(0.126)	(0.317)	(0.166)

***,**,* indicate significance at 1%, 5% and 10% respectively. P-values are in parenthesis. The complete regression results are shown in Appendix A. 7

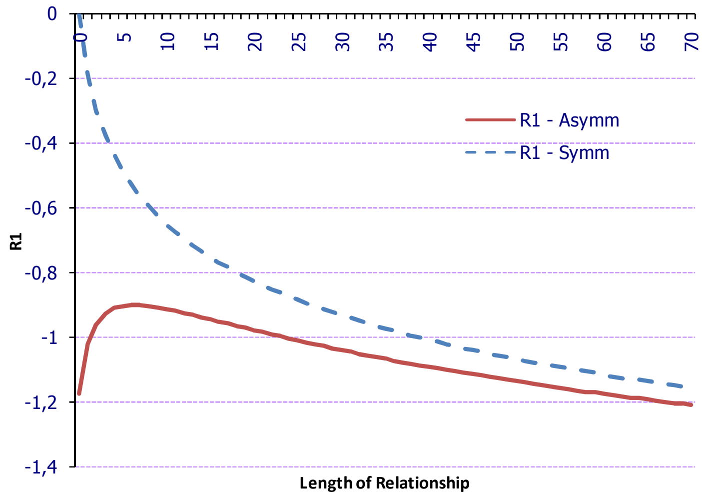


Figure 26 Predicted evolution of interest rates according to the estimation results in the "originally owned firms" sample

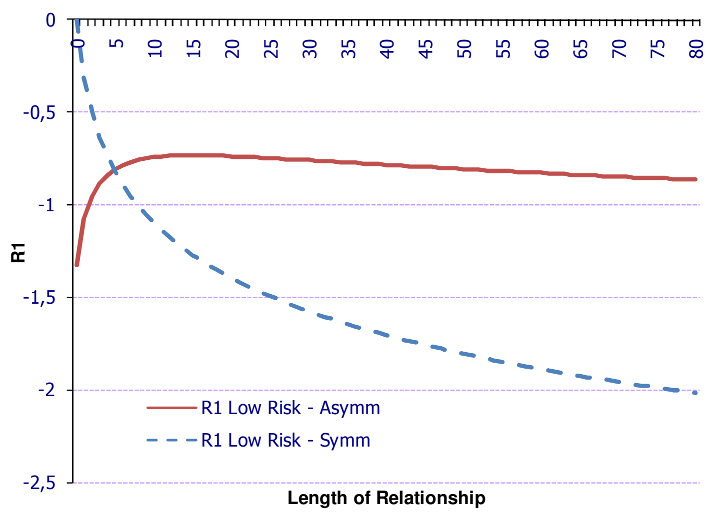


Figure 27 Predicted evolution of interest rates according to the estimation results in the sample of firms with the highest levels of internal funds (Low Risk sample)

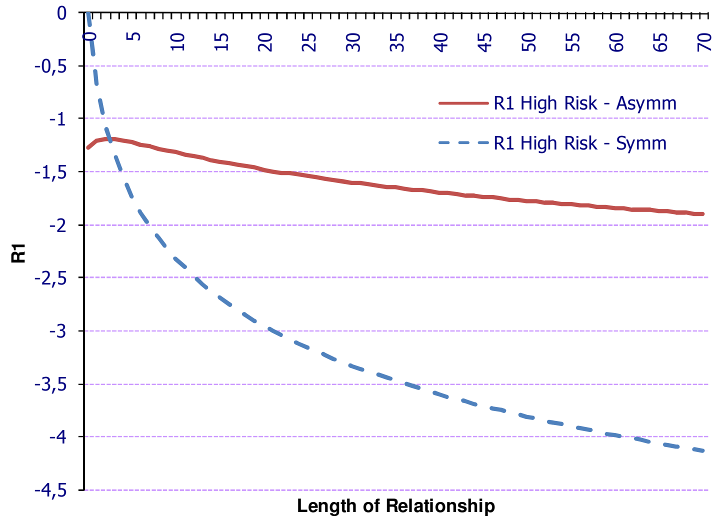


Figure 28 Predicted evolution of interest rates according to the estimation results in the sample of firms with the lowest levels of internal funds (High Risk sample)

5.3.3 Other Robustness Checks

In order to check the robustness of the results, I run the same regressions estimated in the previous sections but using different functional specifications. In the new specifications I have changed the log of the variables length and age, by their square roots, by the variables in levels, or by a polynomial specification that includes both the variables in levels and its square root (or the square of the variable).

In Table 36 and Table 37 we can see exactly the same regression results shown in Table 34 and Table 35, but using the square root of LENGTHM and AGE instead of the logarithm of these variables. In Figure 29, Figure 30 and Figure 31 we can also see the implied evolution of interest rates using this specification. We can see that in this case the estimated results are even closer to the results of the simulated sample if we compared them to Figure 23 to 25.

Although not shown here, the results of the regression under a polynomial specification are very similar to the ones including the square root. In the case of including the variables LENGTHM and AGE in levels, it is of course not possible to find a non-linear evolution of interest rates. However, the sign of the variables and the significance are perfectly in line with the theoretical framework. For instance, in the "Asymmetric" sample the coefficient of the length is positive and the coefficient of the age is negative, but the coefficient of the length is higher than the age in absolute terms, which implies that interest rates increase after the start of a relationship. Meanwhile, in the "symmetric" case the coefficient of the length is negative and significant. Similar results are obtain in the Low Risk versus High Risk regressions.

It is also important to notice that in all the specifications, in order to obtain a significant effect of the variable length it is necessary to use only the sample of firms established by their current owners and without outliers (LENGTHM/AGE<5 or smaller).

Table 36 Basic regression and Split Sample Regressions using the Square Root

	Whole Sample	Original Owned	Low risk	High risk	Asymmetric	Symmetric
	Number of obs= 1672	Number of obs = 1043	Number of obs = 626	Number of obs = 417	Number of obs = 533	Number of obs = 469
	F(60, 1611)= 9.37	F(60, 982) = 6.92	F(60, 565) = 7.33	F(60, 356) = 2.93	F(60, 472) = 5.90	F(60, 408) = 4.29
	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000
	R-squared = 0.2620	R-squared = 0.2767	R-squared = 0.3213	R-squared = 0.3580	R-squared = 0.3504	R-squared = 0.3406
SQR(LENGTHM)	0.0241	0.1912**	0.2631***	0.1179	0.3065***	0.5777
	(0.618)	(0.021)	(0.008)	(0.481)	(0.010)	(0.265)
SQR(AGE)	-0.1450**	-0.2932***	-0.2402**	-0.3552**	-0.2886**	-0.8374*
	(0.013)	(0.002)	(0.035)	(0.046)	(0.019)	(0.088)
LN(AMOUNT)	-0.3081***	-0.4058***	-0.4496***	-0.2942**	-0.3569***	-0.4616***
	(0.000)	(0.000)	(0.000)	(0.013)	(0.000)	(0.000)
LN(ASSETS)	-0.0999*	-0.0320	-0.0306	-0.0859	-0.0734	0.0051
	(0.053)	(0.646)	(0.715)	(0.477)	(0.328)	(0.966)

Table 37 More restricted samples using the Square Root

	Low Risk		High Risk	
	Number of obs = 334	Number of obs = 249	Number of obs = 199	Number of obs = 186
	F(59, 273) = .	F(58, 190) = 5.61	F(59, 138) = .	F(58, 127) = 4.24
	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000
	R-squared = 0.4139	R-squared = 0.4157	R-squared = 0.4720	R-squared = 0.5553
	Asymmetric	**Symmetric**	**Asymmetric**	**Symmetric**
SQR(LENGTHM)	0.3271**	-0.0404	0.2685	-0.6670***
	(0.018)	(0.838)	(0.321)	(0.002)
SQR(AGE)	-0.2756**		-0.4109	
	(0.031)		(0.172)	
LN(AMOUNT)	-0.4236***	-0.4784***	-0.1530	-0.4665**
	(0.004)	(0.008)	(0.326)	(0.015)
LN(ASSETS)	-0.0594	-0.1771	-0.2234	0.0214
	(0.541)	(0.315)	(0.103)	(0.906)

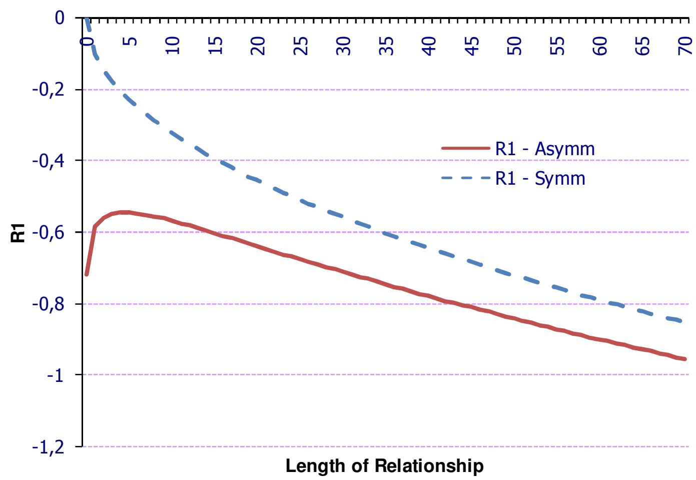


Figure 29 Predicted evolution of interest rates in the "originally owned firms" sample using the square root

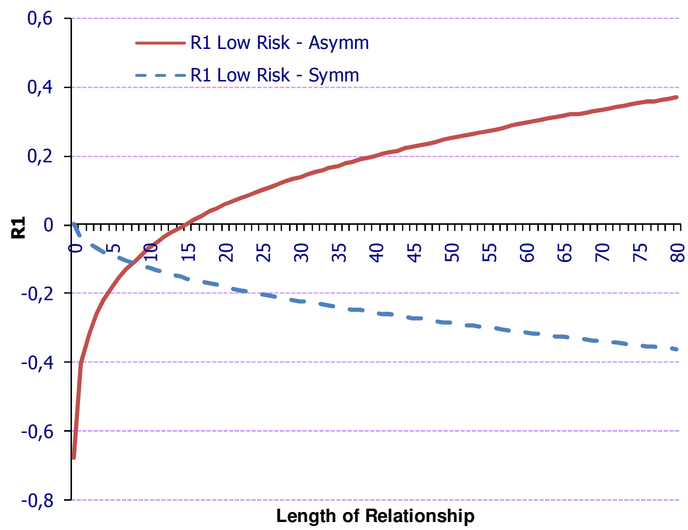


Figure 30 Predicted evolution of interest rates in the "Low Risk" sample using the square root

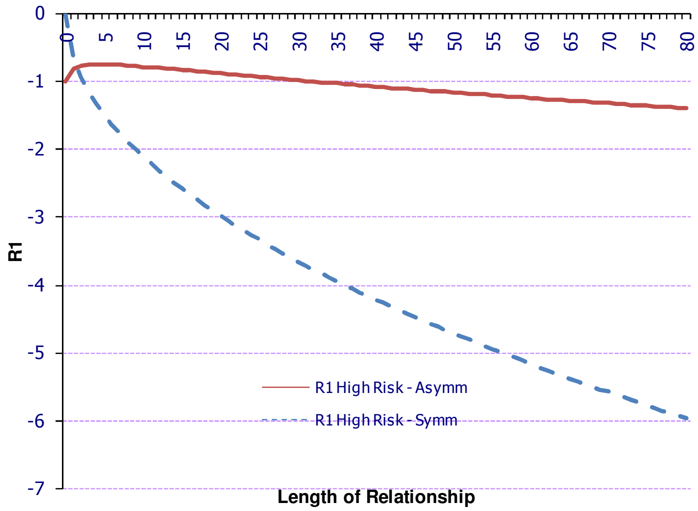


Figure 31 Predicted evolution of interest rates in the "High Risk" sample using the square root

5.4 Conclusions

The theoretical model presented in Chapter 4 and the theoretical propositions of this chapter suggest that interest rates may increase with a bank's amount of information (the length of a relationship) when there is an asymmetry of information among the bank and the firm. This result is found even though the harm of financial frictions diminishes at the same time that interest rates increase.

However, the theoretical model implies that the final effect of the length of relationship depends on degree of asymmetry of information and on the level of internal funds (leverage) of every firm. With symmetry of information we should find that interest rates decrease monotonically as the length increase, but with asymmetry we should see that the effect may depend on the level of risk (internal funds).

On average for all the firms, we expect that interest rates under asymmetry of information increase after the start of a relationship, but that they start decreasing after some time. Additionally, interest rates should decrease or do not change after the start of a relationship in the asymmetric case for the riskier firms. And we should find that interest rates increase or show a hump shaped evolution in the least risky firms' case.

This implication is independent of any effect that a relationship could have on interest rates due to lock in or hold up problems. Actually, the novelty of the results is not that interest rates could increase with the length of a relationship but that they could increase even if we ignore those problems mentioned before. Another novel implication is that the effect of the length on interest rates could be different depending on the relative amount of information that banks and firms have.

I test these predictions using the National Survey of Small Business Finance 2003 and I find evidence supporting this type of evolution. First, I find that the estimated coefficients are consistent with a hump shaped evolution of interest rates through the course of a relationship. However, if I restrict the sample to those firms with higher levels of leverage (higher risk) the estimated coefficients indicate a decreasing evolution of interest rates. Similarly, if I restrict the sample to those firms with medium and low levels of leverage, the estimated coefficients again support a hump shaped evolution. Furthermore, the results confirm that such hump shaped evolution is more likely to occur with asymmetry of information and that on the contrary, longer relationships reduce interest rates when there is symmetry of information.

These results support the theoretical implications of the model. They can be differentiated from alternative explanations such as those coming from lock in or hold up problems because of two main reasons: First of all, models based on the lock-in problem do not predict a different evolution of interest rates depending on whether the length of relationship is equal or smaller than the age of the firm, meanwhile the model presented here predicts that interest rates decrease with longer relationships when the length is equal to the age of the firm and a hump-shape evolution when they are different; something that is confirmed with the data. Secondly, in a lock-in problem model the evolution of interest rates should be the same independently of the level of risk of the firm (more or less leverage or internal funds), meanwhile the empirical evidence presented herein shows that the effect is indeed different.

The results of the chapter are also important because they could help to clarify why the empirical results of the different studies that have studied the relationship between interest rates on commercial loans and the length of a relationship using similar US data is mixed and inconclusive. The main added value of the empirical analysis is to show that the effect of the length of relationship on loan characteristics can only be properly understood by considering the role of internal funds and the perception of risk on the demand for debt, and the role of the length relative to the age of the firm.

Appendix A. 1

CHAPTER 3 VARIABLE DEFINITIONS AND SOURCES

Variable	Definition	Source
HINT	Perception on how problematic is for the firm the high interest rates	World Business Environmental Survey (WBES) 2000. World Bank.
LTLOAN	Perception on how problematic is for the firm the access to long term loans	World Business Environmental Survey (WBES) 2000. World Bank.
COLL	Perception on how problematic is for the firm the collateral requirements of banks/financial institutions.	World Business Environmental Survey (WBES) 2000. World Bank.
PAPR	Perception on how problematic is for the firm the bank paperwork/bureaucracy	World Business Environmental Survey (WBES) 2000. World Bank.
LCKM	Perception on how problematic is for the firm the banks lack money to lend	World Business Environmental Survey (WBES) 2000. World Bank.
SPCN	Perception on how problematic is for the firm the need special connections with banks/financial institutions	World Business Environmental Survey (WBES) 2000. World Bank.
ACNB	Perception on how problematic is for the firm the access to non bank equity/investors/partners	World Business Environmental Survey (WBES) 2000. World Bank.
CRD	Perception on how problematic is for the firm the inadequate credit information on customers	World Business Environmental Survey (WBES) 2000. World Bank.
GCF	Perception on how problematic is for the firm the general constraint finance	World Business Environmental Survey (WBES) 2000. World Bank.
IAS	Dummy variable that takes the value of 1 if the firm is using international financial statements	World Business Environmental Survey (WBES) 2000. World Bank.
AFS	Dummy variable that takes the value of 2 if the firm has audited financial statements and the value of 1 if not.	World Business Environmental Survey (WBES) 2000. World Bank.
SIZE_F	Categorical variable measuring the size of the firm according to the number of employees. It takes the value of 1 for small firms, 2 for medium size firms and 3 for large firms.	World Business Environmental Survey (WBES) 2000. World Bank.
AGE	Number of years since the firm was originally established.	World Business Environmental Survey (WBES) 2000. World Bank.
BANK	Proportion of investment in the past three years financed through domestic and foreign commercial banks debt.	World Business Environmental Survey (WBES) 2000. World Bank.
FN_RE	Proportion of investment in the past three years financed through internal funds/Retained earnings	World Business Environmental Survey (WBES) 2000. World

		Bank.
FN_SHARE	Proportion of investment in the past three years financed through equity, sale of stock	World Business Environmental Survey (WBES) 2000. World Bank.
INFORMAL	Proportion of investment in the past three years financed through informal sources of finance (moneylenders, family or friends).	World Business Environmental Survey (WBES) 2000. World Bank.
FN_SCCR	Proportion of investment in the past three years financed through supplier credit.	World Business Environmental Survey (WBES) 2000. World Bank.
BNK	Dummy variable that takes the value of 1 if the firm is using bank credit.	World Business Environmental Survey (WBES) 2000. World Bank.
SCCR_D	Dummy variable that takes the value of 1 if the firm is using supplier credit.	World Business Environmental Survey (WBES) 2000. World Bank.
INFORMAL_D	Dummy variable that takes the value of 1 if the firm is using informal sources.	World Business Environmental Survey (WBES) 2000. World Bank.
INVESTM	Rate of growth of Investment in the last three years	World Business Environmental Survey (WBES) 2000. World Bank.
SALES_F	Predicted Rate of growth of Future Sales in the next three years	World Business Environmental Survey (WBES) 2000. World Bank.
SALES	Rate of growth of sales in the last three years	World Business Environmental Survey (WBES) 2000. World Bank.
DEBT	Rate of growth of debt in the last three years	World Business Environmental Survey (WBES) 2000. World Bank.
GROWTH	Growth rate of GDP, average 1995-99	United Nations Common Database
CONCENTR	Assets of three largest banks as share of all commercial banks in the system.	World Bank Databases. Original Source: Fitch's Bankscope Database
LEGAL_RIGHT	Index of the strength of legal rights. It measures the degree to which collateral and bankruptcy laws protect the rights of borrowers and lenders and thus facilitate lending. The index includes 7 aspects related to legal rights in collateral law and 3 aspects in bankruptcy law.	"Doing Business" World Bank Database 2004. These variables were constructed by Djankov, McLiesh and Shleifer (forthcoming).
CRED_INFORM	index of the depth of credit information and it measures rules affecting the scope, accessibility and quality of credit information available through either public or private credit registries	"Doing Business" World Bank Database 2004. These variables were constructed by Djankov, McLiesh and Shleifer (forthcoming).
PUB_REGSTR	indicator of the public credit registry coverage. It reports the number of individuals and firms listed in a public credit registry with current information on repayment history, unpaid debts or credit outstanding	"Doing Business" World Bank Database 2004. These variables were constructed by Djankov, McLiesh and Shleifer (forthcoming).
PRIV_BUREAU	indicator of the private credit bureau coverage and reports the number of individuals or firms listed by a	"Doing Business" World Bank Database 2004. These variables

	private credit bureau with current information on repayment history, unpaid debts or credit outstanding	were constructed by Djankov, McLiesh and Shleifer (forthcoming).

Appendix A. 2

Derivation of the prior distribution

The prior parameters can be simply assumed as coming from a subjective belief of the entrepreneur about the distribution of θ.

In a different way, I assume that these values are not completely subjective, but that the firm may use some other source of information to determine its prior beliefs. In the case proposed here, an entrepreneur starting a new firm may have also learned about his own ability through working experience, and may use as prior for the firm's quality, the average quality of a firm in the industry (sector) from which the firm belongs.

Similar to Jovanovic (1982) I assume that the entrepreneur's ability θ_m is a random draw from the population's ability which is distributed $N(\theta_{mP} - \sigma_{mP}^2, \sigma_{mP}^2)$. So when an entrepreneur does not have any working experience his initial belief is that his own ability is equal to the average ability in the population. After working for H periods the entrepreneur updates his beliefs based on the H realizations of his own productivity m_t. Therefore, before the start-up the prior distribution for the entrepreneur's ability θ_m has a mean θ_{0m} given by:

(A.2. 1) $$\theta_{0m} = E_0(\theta_m \mid \Omega_F, T=0) = \frac{(\hat{\theta}_m)(\sigma_m^2 / H)^{-1} + (\theta_{mP})(\sigma_{mP}^2)^{-1}}{(\sigma_m^2 / H)^{-1} + (\sigma_{mP}{}^2)^{-1}}$$

And a variance σ_{0m}^2 given by:

(A.2. 2) $$\sigma_{0m}^2 = \operatorname{var}(\theta_m \mid \Omega_F, T=0) = \left(\frac{1}{(\sigma_m^2 / H)^{-1} + (\sigma_{mP}^2)^{-1}} \right)$$

Where $\hat{\theta}_m$ is the estimated ability during the H periods of working experience. The prior variance decreases with higher experience of the entrepreneur (higher H). This means that entrepreneurs with more experience should have a better estimation of their own ability before the start-up. Similarly, I assume that the firm's true quality θ_z is also a random draw from the quality in the industry, which is distributed $N(\theta_{zI} - \sigma_{zI}^2 / 2, \sigma_{zI}^2)$. So when an entrepreneur starts a new firm his prior belief about the firm's quality should be just equal to the average quality in the industry θ_{zI}. Therefore, before the start-up the prior distribution of the firm quality θ_z has a mean θ_{0z} given by:

(A.2. 3) $$\theta_{0z} = E_0(\theta_z \mid \Omega_F, T=0) = \theta_{zI}$$

And a variance σ_{0z}^2 given by:

(A.2. 4) $$\sigma_{0z}^2 = \operatorname{var}(\theta_z \mid \Omega_F, T=0) = \sigma_{zI}^2$$

Based on the prior distributions for his ability and the firm's quality, the entrepreneur can construct a prior distribution for the parameter θ, and this prior distribution has a mean θ_{0F} given by:

(A.2. 5) $$\theta_{0F} = E_0(\theta \mid \Omega_F, T=0) = E_0(\theta_m + \theta_z \mid \Omega_F, T=0) = \theta_{0m} + \theta_{zI}$$

And a variance σ_{0F}^2 given by:

(A.2. 6) $$\sigma_{0F}^2 = \text{var}(\theta \mid \Omega_F, T=0) = \text{var}(\theta_m + \theta_z \mid \Omega_F, T=0) = \sigma_{0m}^2 + \sigma_{zI}^2$$

With $\theta_{0m}, \sigma_{0m}^2$ as defined in equations (9) and (10) respectively. The most important issue to notice about the prior distribution described here is that entrepreneurs may learn something before they start a new firm and therefore, entrepreneurs with higher experience should have more precise estimations of their future profits even before the start-up. However, the true distribution of θ_t or the true profitability of the firm can only be learned after production starts, because the true quality of the firm cannot be learned before the start-up.

In order to understand how the different functions related with the variance of prediction evolve as both firm and bank learn, I introduce the following formal definitions:

Definition A. 1 The cumulative distribution function for the prediction Θ_1 is given by: $\Phi_F(\Theta_1) = \frac{1}{2}\left(1 + erf\left(\frac{\theta_1 - \theta_F + \sigma_F^2/2}{\sigma_F\sqrt{2}}\right)\right)$. Where $\ln(\Theta_1) = \theta_1$ and *erf* is called the error function.

Definition A. 2 The firm's partial expectations of Θ_1 above and below the default point $\overline{\Theta}_1$ are denoted as $(\tilde{\Theta}_{1F} \mid S)$ and $(\tilde{\Theta}_{1F} \mid D)$ respectively, and are given by:

$$(\tilde{\Theta}_{1F} \mid S) = \exp(\theta_{1F})\left(1 - \Phi_{SN}\left((\overline{\theta}_1 - \theta_F - \sigma_F^2/2)/\sigma_F\right)\right)$$

$$(\tilde{\Theta}_{1F} \mid D) = \exp(\theta_{1F})\left(\Phi_{SN}\left((\overline{\theta}_1 - \theta_F - \sigma_F^2/2)/\sigma_F\right)\right)$$

where Φ_{SN} is the cumulative distribution function of the standard normal.

Definition A. 3 The bank's partial expectations of Θ_1 above and below the default point $\overline{\Theta}_1$ are denoted as $(\tilde{\Theta}_{1B} \mid S)$ and $(\tilde{\Theta}_{1B} \mid D)$ respectively, and are given by:

$$(\tilde{\Theta}_{1B} \mid S) = \exp(\theta_{1B})\left(1 - \Phi_{SN}\left((\overline{\theta}_1 - \theta_B - \sigma_B^2/2)/\sigma_B\right)\right)$$

$$(\tilde{\Theta}_{1B} \mid D) = \exp(\theta_{1B})\left(\Phi_{SN}\left((\overline{\theta}_1 - \theta_B - \sigma_B^2/2)/\sigma_B\right)\right)$$

All the functions defined above change as the firm and the bank learn about the true profitability parameter θ. It is important to understand how these functions change with the variance because that change has two different effects on the expected value of the firm.

Appendix A. 3

PROOFS OF THEORETICAL PROPOSITIONS AND COROLLARIES

Proof of Proposition 1

The default point $\overline{\Theta}_1$ is implicitly defined by equation (17)

$$\overline{\Theta}_1 k_1^{\beta} + \frac{qk_1 - RB_1}{(1-\Phi_B)} - \frac{\tau q k_1}{(1-\Phi_B)} + \frac{(\tilde{\Theta}_{1B} \mid D) k_1^{\beta}}{(1-\Phi_B)} = 0$$

The effect of a change in k_1 in the default point would be given by:

(A.3. 1) $$\frac{\partial \overline{\Theta}_1}{\partial k_1} = -\frac{\frac{\partial F}{\partial k_1}(k_1, \overline{\Theta}_1)}{\frac{\partial F}{\partial \overline{\Theta}_1}(k_1, \overline{\Theta}_1)}$$

where $F(k_1, \overline{\Theta}_1) = \overline{\Theta}_1 k_1^{\beta} + \frac{qk_1 - RB_1}{(1-\Phi_B)} - \frac{\tau q k_1}{(1-\Phi_B)} + \frac{(\tilde{\Theta}_{1B} \mid D) k_1^{\beta}}{(1-\Phi_B)} = 0$, which can be written as $F(k_1, \overline{\Theta}_1) = \left(\overline{\Theta}_1 (1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)\right) k_1^{\beta} - (R - q(1-\tau)) k_1 + RW_0 = 0$.

The partial derivatives of $F(k_1, \overline{\Theta}_1)$ with respect to k_1 and $\overline{\Theta}_1$ are given by:

(A.3. 2) $$\frac{\partial F}{\partial k_1} = \beta k_1^{\beta-1} \left(\overline{\Theta}_1 (1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)\right) - (R - q(1-\tau))$$

(A.3. 3) $$\frac{\partial F}{\partial \overline{\Theta}_1} = \left((1-\Phi_B) - \overline{\Theta}_1 \frac{\partial \Phi_B}{\partial \overline{\Theta}_1} + \frac{\partial (\tilde{\Theta}_{1B} \mid D)}{\partial \overline{\Theta}_1} \right) k_1^{\beta}$$

Thus, equation (A.3. 1) is equal to:

(A.3. 4) $$\frac{\partial \overline{\Theta}_1}{\partial k_1} = -\frac{\beta k_1^{\beta-1} \left(\overline{\Theta}_1 (1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)\right) - (R-q) - \tau q}{k_1^{\beta} ((1-\Phi_B) - \overline{\Theta}_1 \frac{\partial \Phi_B}{\partial \overline{\Theta}_1} + \frac{\partial (\tilde{\Theta}_{1B} \mid D)}{\partial \overline{\Theta}_1})}$$

In order to solve for this expression we need to derive the cumulative distribution function Φ_B and the partial expectation $(\tilde{\Theta}_{1B} \mid D)$ with respect to $\overline{\Theta}_1$.

Using Definition A. 2 and Definition A. 3 from Appendix A. 2, and deriving them with respect to $\overline{\Theta}_1$, we have that

$$\frac{\partial \Phi_B}{\partial \overline{\Theta}_1} = \phi_B = \frac{\exp(-z_1^2)}{\overline{\Theta}_1 \sigma_B \sqrt{2\pi}} > 0$$

And $\frac{\partial(\tilde{\Theta}_{1B} \mid D)}{\partial\overline{\Theta}_1} = -\frac{\tilde{\Theta}_{1B}\exp(-z_2^2)}{\sqrt{2\pi}}\left(-\frac{1}{\sigma_B\overline{\Theta}_1}\right) > 0$

With $z_2 = \frac{(-\overline{\theta}_1 + \theta_B + \sigma_B^2/2)}{\sigma_B\sqrt{2}}$, $z_1 = \left(\frac{\overline{\theta}_1 - \theta_B + \sigma_B^2/2}{\sigma_B\sqrt{2}}\right)$ and $\tilde{\Theta}_{1B} = \exp(\theta_B)$.

It can be shown that $z_2^2 = z_1^2 - \overline{\theta}_1 + \theta_B$. Using this last expression, we should have that $\tilde{\Theta}_{1B}\exp(-z_2^2) = \exp(\theta_B)\exp(-z_1^2 + \overline{\theta}_1 - \theta_B) = \exp(-z_1^2 + \overline{\theta}_1) = \exp(-z_1^2)\exp(\overline{\theta}_1)$ or that

(A.3. 5) $$\tilde{\Theta}_{1B}\exp(-z_2^2) = \exp(\theta_B)\exp(-z_2^2) = \exp(-z_1^2)\overline{\Theta}_1$$

Therefore we have that $\frac{\partial(\tilde{\Theta}_{1B} \mid D)}{\partial\overline{\Theta}_1} = -\frac{\exp(-z_1^2)\overline{\Theta}_1}{\sqrt{2\pi}}\left(-\frac{1}{\sigma_B\overline{\Theta}_1}\right) = \phi_B\overline{\Theta}_1$

Using this result on equation (A.3. 3) we have that

(A.3. 6) $$\frac{\partial F}{\partial\overline{\Theta}_1}(k_1, \overline{\Theta}_1) = (1-\Phi_B)k_1^{\beta} - \overline{\Theta}_1\phi_B k_1^{\beta} + \overline{\Theta}_1\phi_B k_1^{\beta} = (1-\Phi_B)k_1^{\beta} > 0$$

Finally, we can write the derivative of $\overline{\Theta}_1$ with respect to k_1 as the following:

(A.3. 7) $$\frac{\partial\overline{\Theta}_1}{\partial k_1} = \frac{(R - q(1-\tau)) - \beta k_1^{\beta-1}\left(\overline{\Theta}_1(1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)\right)}{k_1^{\beta}(1-\Phi_B)}$$

The conditions for this derivative to exist are given by the implicit function theorem. I first consider the point $(k_1, \overline{\Theta}_1) = (k_1^0, 0)$, where $k_1^0 = \frac{RW_0}{R - q(1-\tau)}$ represents the threshold above the one the demand for capital k_1 makes that debt B_1 becomes risky. At this point we have that $F(k_1^0, 0) = 0$. Given also that Φ_B is continuous differentiable on the domain $(0,\infty)$ that contains the point $\overline{\Theta}_1 = 0$, that $f(k_1)$ is continuous differentiable on the domain $(0,\infty)$ that contains k_1^0, we should have that $F(k_1, \overline{\Theta}_1)$ is continuous and first differentiable on an open disk with center $(k_1^0, 0)$. Provided also that $\frac{\partial F}{\partial\overline{\Theta}_1}(k_1^0, 0) = (k_1^0)^{\beta}(1-\Phi(0)) = (k_1^0)^{\beta} \neq 0$, we have that:

$\frac{\partial\overline{\Theta}_1}{\partial k_1} = \frac{\frac{\partial F}{\partial k_1}(k_1^0, 0)}{\frac{\partial F}{\partial\overline{\varepsilon}_1}(k_1^0, 0)} = -\frac{-(R - q(1-\tau))}{(k_1^0)^{\beta}} > 0$ is the derivative of a function $\overline{\Theta}_1 = \varphi(k_1)$ on a neighborhood of the point $(k_1^0, 0)$, which means that the default point $\overline{\Theta}_1$ increases as k_1 increases around the point where debt becomes risky, and therefore, the probability of default $\Phi(\overline{\Theta}_1)$ also increases with k_1 when $k_1 > k_1^0$ for k_1 in a neighborhood of k_1^0.

Here, it is important to remember that according to equation (A.3. 7), $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ should be positive whenever its numerator is positive, i.e. whenever $R - q(1-\tau) > \left(\overline{\Theta}_1(1-\Phi_F) + (\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{\beta-1}$.

To complete the proof for every k_1, we first need to prove that the term $\left(\overline{\Theta}_1(1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)\right)$ is always smaller or equal than $\tilde{\Theta}_{1B}$. This can be proved by replacing equation (15), which says that $\overline{\Theta}_1 k_1^{\beta} - \tau q k_1 = R_1 B_1 - q k_1$, into equation (18) and expressing the expected value of the firm (for the bank) as $E_0(V_0) = \frac{1}{R}\left[k_1^{\beta}\left((\tilde{\Theta}_{1B} \mid S) - (1-\Phi_B)\overline{\Theta}_1\right) + (1-\Phi_B)\tau q k_1\right]$. Thus, for the expected value of the firm to be positive it is necessary that $(\tilde{\Theta}_{1B} \mid S) \geq (1-\Phi_B)\overline{\Theta}_1$. Adding $(\tilde{\Theta}_{1B} \mid D)$ to both sides of this latter inequality we have that $(\tilde{\Theta}_{1B} \mid S) + (\tilde{\Theta}_{1B} \mid D) \geq (1-\Phi_B)\overline{\Theta}_1 + (\tilde{\Theta}_{1B} \mid D)$, and thus we have that $\tilde{\Theta}_{1B} \geq \overline{\Theta}_1(1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)$.[22]

Now we can consider the point k_1^* which corresponds to the "first-best" of the firm according to bank's information, i.e. at k_1^* corresponds to the level of capital such that $\beta\tilde{\Theta}_{1B} k_1^{*\beta-1} = R - q$. We can denote the default point associated to k_1^* as $\overline{\Theta}_1^*$. Since we have seen that $\tilde{\Theta}_{1B} > \overline{\Theta}_1^*(1-\Phi_F) + (\tilde{\Theta}_{1B} \mid D)$ we have that at the point $\left(k_1^*, \overline{\Theta}_1^*\right)$, $R - q(1-\tau) > R - q = \beta\tilde{\Theta}_{1B} k_1^{*\beta-1}$ and therefore $R - q(1-\tau) > \left(\overline{\Theta}_1^*(1-\Phi_F) + (\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{*\beta-1}$. Thus, around the point $\left(k_1^*, \overline{\Theta}_1^*\right)$ the numerator of $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ is also positive ant this implies that $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ >0.

Summarizing, we now know that the default point $\overline{\Theta}_1 \in (0, \overline{\Theta}_1^*)$ for any $k_1 \in (k_1^0, k_1^*)$. This implies that $\left(\overline{\Theta}_1(1-\Phi_F) + (\tilde{\Theta}_{1B} \mid D)\right) \in \left(0, \overline{\Theta}_1^*(1-\Phi_F) + (\tilde{\Theta}_{1B} \mid D)\right)$ for any $k_1 \in (k_1^0, k_1^*)$.

Since the function $\overline{\Theta}_1(1-\Phi_F) + (\tilde{\Theta}_{1B} \mid D)$ is strictly increasing in $\overline{\Theta}_1$ (its derivative with respect to $\overline{\Theta}_1$ is equal to $(1-\Phi_F)$) we have that $\overline{\Theta}_1(1-\Phi_F) + (\tilde{\Theta}_{1B} \mid D)$ should be strictly increasing for any point $(k_1, \overline{\Theta}_1)$ between $(k_1^0, 0)$ and $\left(k_1^*, \overline{\Theta}_1^*\right)$. Since at k_1^* we have that $R - q(1-\tau) > \left(\overline{\Theta}_1^*(1-\Phi_F) + (\tilde{\Theta}_{1B} \mid D)\right)\beta \overline{k}_1^{\beta-1}$ and since $R - q(1-\tau)$ is a constant, therefore $\left(\overline{\Theta}_1(1-\Phi_F) + (\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{\beta-1} < R - q(1-\tau)$ for any $k_1 \in (k_1^0, k_1^*)$.

---

[22] This can also be shown in statistical terms by using that $(\tilde{\Theta}_{1B} \mid D) = \tilde{\Theta}_{1B}\Phi_B\left((\exp(\overline{\theta}_1 - \sigma_B/2)\right)$ and that $\Phi_B[(\exp(\overline{\theta}_1 - \sigma_B/2)] < \Phi_B$.

Finally, for any $k_1 > k_1^*$ we know that $\beta\tilde{\Theta}_{1B} k_1^{*\beta-1} < R - q$. This should imply that $\left(\overline{\Theta}_1(1-\Phi_F) + (\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{\beta-1} \leq \beta\tilde{\Theta}_{1B} k_1^{*\beta-1} < R - q < R - q(1-\tau)$. Therefore, we have that for any $k_1 > k_1^*$ the numerator of $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ should also be positive ant thus $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ should be positive as well.

This guarantees that $\frac{\partial \overline{\Theta}_1}{\partial k_1}$ >0 for any $k_1 \geq k_1^0$. Since $k_1^0 = \frac{RW_0}{R - q(1-\tau)}$ this is equivalent to showing that the default point is increasing in the level of capital as long as $W_0 \geq 0$, i.e. when the firm has a positive initial value of equity.

Proof of Proposition 2

Using (A.3. 6) we know that $\frac{\partial F}{\partial \overline{\Theta}_1} = (1-\Phi_B) k_1^{\beta} > 0$ and thus we have the following results:

$$\frac{\partial \overline{\Theta}_1}{\partial W_0} = -\frac{\partial F / \partial W_0}{\partial F / \partial \overline{\Theta}_1} = -\frac{R}{k_1^{\beta}(1-\Phi_B)} < 0$$

$$\frac{\partial \overline{\Theta}_1}{\partial \tau} = -\frac{\partial F / \partial \tau}{\partial F / \partial \overline{\Theta}_1} = -\frac{-qk_1}{k_1^{\beta}(1-\Phi_B)} > 0$$

For proving that $\frac{\partial \overline{\Theta}_1}{\partial \theta_B} < 0$, we need to first find the partial derivative $\frac{\partial F}{\partial \theta_B}$, which is equal to:

$$\frac{\partial F}{\partial \theta} = \overline{\Theta}_1 k_1^{\beta} \frac{\partial (1-\Phi_B)}{\partial \theta_B} + k_1^{\beta} \frac{\partial (\tilde{\Theta}_{1B} \mid D)}{\partial \theta_B}.$$

Deriving the cumulative distribution function and the partial expectation given default with respect to θ_B we have that:

$$\frac{\partial (1-\Phi_B)}{\partial \theta_B} = \frac{\exp(-z_1^2)}{\sigma_B \sqrt{2\pi}} = \overline{\Theta}_1 \phi_B$$

$$\frac{\partial (\tilde{\Theta}_{1B} \mid D)}{\partial \theta_B} = \exp(\theta_B + \sigma_B^2/2)(1-\Phi_{SN}(z_2\sqrt{2})) - \frac{\exp(\theta_B + \sigma_B^2/2)\exp(-z_2^2)}{\sigma_B \sqrt{2\pi}}$$

Which can be expressed as:

$$\frac{\partial (\tilde{\Theta}_{1B} \mid D)}{\partial \theta_B} = (\tilde{\Theta}_{1B} \mid D) - \frac{\exp(-z_1^2)\overline{\Theta}_1}{\sigma_B \sqrt{2\pi}} = (\tilde{\Theta}_{1B} \mid D) - \phi_B (\overline{\Theta}_1)^2$$

Therefore, $\frac{\partial F}{\partial \theta_B} = \overline{\Theta}_1 k_1^{\beta} \overline{\Theta}_1 \phi_B + k_1^{\beta}\left((\tilde{\Theta}_{1B} \mid D) - \phi_B (\overline{\Theta}_1)^2\right) = k_1^{\beta}(\tilde{\Theta}_{1B} \mid D) > 0$

And finally,

$$\frac{\partial \overline{\Theta}_1}{\partial \theta_B} = -\frac{\partial F / \partial \theta}{\partial F / \partial \overline{\Theta}_1} = -\frac{(\tilde{\Theta}_{1B} \mid D)}{(1-\Phi_B)} < 0$$

Proof of Corollary 1

The total repayment R_1B_1 is given by:

$$R_1B_1 = (1-\Phi_B)^{-1}\left[RB_1 - \Phi_B qk_1(1-\tau) - (\tilde{\Theta}_{1B} \mid D)k_1^{\beta}\right]$$

Its derivative with respect to k_1 is given by:

$$\frac{\partial R_1B_1}{\partial k_1} = (1-\Phi_B)^{-2}\phi_B \frac{\partial\overline{\Theta}_1}{\partial k_1}\left[RB_1 - \Phi_B qk_1(1-\tau) - (\tilde{\Theta}_{1B} \mid D)k_1^{\beta}\right]$$
$$+(1-\Phi_B)^{-1}\left[R - \phi_B \frac{\partial\overline{\Theta}_1}{\partial k_1}qk_1(1-\tau) - \Phi_B q(1-\tau) - \phi_B\overline{\Theta}_1 \frac{\partial\overline{\Theta}_1}{\partial k_1}k_1^{\beta} - \beta(\tilde{\Theta}_{1B} \mid D)k_1^{\beta-1}\right]$$

Factorizing the terms that contain $\phi_B \frac{\partial\overline{\Theta}_1}{\partial k_1}$ we can write the previous expression as:

$$\frac{\partial R_1B_1}{\partial k_1} = (1-\Phi_B)^{-2}\phi_B \frac{\partial\overline{\Theta}_1}{\partial k_1}\left[RB_1 - qk_1(1-\tau) - \left((\tilde{\Theta}_{1B} \mid D) + \overline{\Theta}_1(1-\Phi_B)\right)k_1^{\beta}\right]$$
$$+(1-\Phi_B)^{-1}\left[R - \Phi_B q(1-\tau) - \beta(\tilde{\Theta}_{1B} \mid D)k_1^{\beta-1}\right]$$

Since the first term in square brackets is equal to zero (because of the equation that defines the default point) we finally have that

(A.3. 8) $$\frac{\partial R_1B_1}{\partial k_1} = (1-\Phi_B)^{-1}\left[R - \Phi_B q(1-\tau) - \beta(\tilde{\Theta}_{1B} \mid D)k_1^{\beta-1}\right] > 0$$

Which is positive because, according to the proof for Proposition 1, $R - q(1-\tau) > \left(\overline{\Theta}_1(1-\Phi_F) + (\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{\beta-1}$ is always positive, or in other words, because $\frac{\partial\overline{\Theta}_1}{\partial k_1}$ >0.

To prove that the total repayment is a strictly decreasing function of W_0, we need to derive the total repayment R_1B_1 with respect to W_0 and we get the following:

$$\frac{\partial R_1B_1}{\partial W_0} = (1-\Phi_B)^{-2}\phi_B \frac{\partial\overline{\Theta}_1}{\partial W_0}\left[RB_1 - \Phi_B qk_1(1-\tau) - (\tilde{\Theta}_{1B} \mid D)k_1^{\beta}\right]$$
$$+(1-\Phi_B)^{-1}\left[-R - \phi_B \frac{\partial\overline{\Theta}_1}{\partial W_0}qk_1(1-\tau) - \phi_B\overline{\Theta}_1 \frac{\partial\overline{\Theta}_1}{\partial W_0}k_1^{\beta}\right]$$

Following the same steps as in the previous case we get:

$$\frac{\partial R_1B_1}{\partial W_0} = \frac{R}{(1-\Phi_B)} > 0$$

For proving that the total repayment is a strictly convex function of k_1 we just derive equation (A.3. 8) with respect to k_1 and we get (factorizing some common terms):

$$\frac{\partial^2 R_1 B_1}{\partial^2 k_1} = \frac{\phi_B \frac{\partial \overline{\Theta}_1}{\partial k_1}\left[R - q(1-\tau) - \beta\left((\tilde{\Theta}_{1B} \mid D) + \overline{\Theta}_1(1-\Phi_B)\right)k_1^{\beta-1}\right]}{(1-\Phi_B)^2} - \frac{\beta(\beta-1)(\tilde{\Theta}_{1B} \mid D)k_1^{\beta-2}}{(1-\Phi_B)} > 0$$

This expression is positive since (β - 1) < 0 and since the term in square brackets is always positive because $\frac{\partial \overline{\Theta}_1}{\partial k_1} > 0$.

Similarly, we have that the loan's interest rate is given by:

$$R_1 = (1-\Phi_B)^{-1} B_1^{-1}\left[RB_1 - \Phi_B q k_1(1-\tau) - (\tilde{\Theta}_{1B} \mid D)k_1^{\beta}\right]$$

Deriving this expression with respect to k_1 we get the following (omitting the algebra which is similar to the previous cases):

(A.3. 9)
$$\frac{\partial R_1}{\partial k_1} = \frac{\left[\Phi_B q k_1(1-\tau)(k_1 - B_1) + (\tilde{\Theta}_{1B} \mid D)k_1^{\beta}(1-\beta) + \beta(\tilde{\Theta}_{1B} \mid D)k_1^{\beta-1}W_0\right]}{(1-\Phi_B)B_1^2} > 0$$

Which is always positive because $k_1 > B_1$ and (1-β) > 0.

Proof of Proposition 3

The value of the firm at t = 1 is equal to:

$$\Theta_1 k_1^{\beta} + q k_1 - R_1 B_1$$

The second derivative of this function with respect to k_1 is equal to:

$$\beta(\beta-1)\Theta_1 k_1^{\beta-2} - \frac{\phi_B \frac{\partial \overline{\Theta}_1}{\partial k_1}\left[R - q(1-\tau) - \beta\left((\tilde{\Theta}_{1B} \mid D) + \overline{\Theta}_1(1-\Phi_B)\right)k_1^{\beta-1}\right]}{(1-\Phi_B)^2} + \frac{\beta(\beta-1)(\tilde{\Theta}_{1B} \mid D)k_1^{\beta-2}}{(1-\Phi_B)} < 0$$

The expression is negative since (β - 1) < 0 and since the term in square brackets is always positive because $\frac{\partial \overline{\Theta}_1}{\partial k_1} > 0$.

Therefore, the value of the firm at t = *1* is a strictly concave function of k_1.

Proof of Proposition 4

This proposition refers to the sign of $\frac{\partial \overline{\Theta}_1}{\partial \sigma_B}$ that is given by:

$$\frac{\partial \overline{\Theta}_1}{\partial \sigma_B} = -\frac{\frac{\partial F}{\partial \sigma_B}(k_1, \overline{\Theta}_1)}{\frac{\partial F}{\partial \overline{\Theta}_1}(k_1, \overline{\Theta}_1)}$$

Using equation (A.3. 6) we know that $\frac{\partial F}{\partial \overline{\Theta}_1} > 0$, and therefore is clear that the sign of $\frac{\partial \overline{\Theta}_1}{\partial \sigma_B}$ depends only on the sign of $\frac{\partial F}{\partial \sigma_B}$ (or the sign of the numerator). When this sign is positive, $\frac{\partial \overline{\Theta}_1}{\partial \sigma_B}$ should be negative and when the sign of $\frac{\partial F}{\partial \sigma_B}$ is negative, then $\frac{\partial \overline{\Theta}_1}{\partial \sigma_B}$ should be positive. Deriving $F(k_1, \overline{\Theta}_1)$ with respect to σ_B we have that:

(A.3. 10) $$\frac{\partial F}{\partial \sigma_B} = \overline{\Theta}_1 \frac{\partial(1-\Phi_B)}{\partial \sigma_B} k_1^{\beta} + \frac{\partial(\tilde{\Theta}_{1B} \mid D)}{\partial \sigma_B} k_1^B = \left(\overline{\Theta}_1 \frac{\partial(1-\Phi_B)}{\partial \sigma_B} + \frac{\partial(\tilde{\Theta}_{1B} \mid D)}{\partial \sigma_B} \right) k_1^B$$

Deriving the cumulative function $\Phi_B(\overline{\Theta}_1)$ with respect to its variance and we obtain:

(A.3. 11) $$\frac{\partial \Phi_B(\overline{\Theta}_1)}{\partial \sigma_B} = -\left(\frac{\exp(-z_1^2)}{\sqrt{2\pi}\sigma_B^2} \left(\overline{\theta}_1 - \theta_B - \sigma_B^2/2\right) \right)$$

with $z_1 = \left(\frac{\overline{\theta}_1 - \theta_B + \sigma_B^2/2}{\sigma_B\sqrt{2}} \right)$

Deriving the partial expectation below the default point $(\tilde{\Theta}_{1B} \mid D)$ with respect to the variance, we get:

(A.3. 12) $$\frac{\partial(\tilde{\Theta}_{1B} \mid D)}{\partial \sigma_B} = -\frac{\exp(-z_2^2)\exp(\theta_B)}{\sqrt{2\pi}\sigma_B^2} \left(\overline{\theta}_1 - \theta_B + \sigma_B^2/2\right)$$

With $z_2 = \frac{(-\overline{\theta}_1 + \theta_B + \sigma_B^2/2)}{\sigma_B\sqrt{2}}$

Using equation (A.3. 11) and (A.3. 12) we can express the derivatives of $(1-\Phi_B)$ and $(\tilde{\Theta}_{1B} \mid D)$ with respect to σ_B as the following:

$$\frac{\partial(1-\Phi_B)}{\partial \sigma_B} = \left(\frac{\exp(-z_1^2)}{\sqrt{2\pi}\sigma_B^2} \left(\overline{\theta}_1 - \theta_B - \sigma_B^2/2\right) \right)$$

$$\frac{\partial(\tilde{\Theta}_{1B} \mid D)}{\partial \sigma_B} = -\frac{\exp(\theta_B)\exp(-z_2^2)}{\sigma_B^2\sqrt{2\pi}} \left(\overline{\theta}_1 - \theta_B + \sigma_B^2/2\right)$$

Using that $z_2^2 = z_1^2 - \overline{\theta}_1 + \theta$, we established in equation (A.3. 5) in the proof of Proposition 1, that $\exp(\theta_B)\exp(-z_2^2) = \overline{\Theta}_1 \exp(-z_1^2)$. Therefore we have that,

$$\frac{\partial(\tilde{\Theta}_{1B} \mid D)}{\partial \sigma_B} = -\frac{\overline{\Theta}_1 \exp(-z_1^2)}{\sigma_B^2\sqrt{2\pi}} \left(\overline{\theta}_1 - \theta_B + \sigma_B^2/2\right)$$

Replacing these expressions into $\frac{\partial F}{\partial \sigma_B}$, we get the following:

$$\frac{\partial F}{\partial \sigma_B} = \left(\overline{\Theta}_1 \frac{\exp(-z_1^2)}{\sqrt{2\pi}\sigma_B^2} \left(\overline{\theta}_1 - \theta_B - \sigma_B^2/2\right) - \frac{\exp(-z_1^2)\overline{\Theta}_1}{\sigma_B^2\sqrt{2\pi}} \left(\overline{\theta}_1 - \theta_B + \sigma_B^2/2\right) \right) k_1^B$$

Which could be simplified into:

(A.3. 13)
$$\frac{\partial F}{\partial \sigma_B} = k_1^B \left(\frac{\exp(-z_1^2)\overline{\Theta}_1}{\sigma_B^2 \sqrt{2\pi}} \left(-\sigma_B^2\right) \right) = -k_1^B \left(\phi_B \overline{\Theta}_1 \sigma_B\right) < 0$$

Therefore (using (A.3. 6) the effect of a change in the variance in the default point is given by:

(A.3. 14)
$$\frac{\partial \overline{\Theta}_1}{\partial \sigma_B} = -\frac{\frac{\partial F}{\partial \sigma_B}}{\frac{\partial F}{\partial \overline{\Theta}_1}} = \frac{\left(\phi_B \overline{\Theta}_1 \sigma_B\right)}{(1-\Phi_B)} > 0$$

Proof of Proposition 5

To prove that the level of investment in the asymmetric case is always smaller than in the symmetric case we have to determine the sign of equation (25), i.e. we need to prove that the following expression is negative:

$$\left(\frac{(1-\Phi_F^{EVD})}{(1-\Phi_F)} \left[(1-\Phi_B)-(1-\Phi_F)\right] + (\Phi_B^{EVD} - \Phi_F^{EVD}) \right) = 0$$

Multiplying both sides of the equation by $(1-\Phi_F)$ we get the following:

$$\left((1-\Phi_F^{EVD})\left[(1-\Phi_B)-(1-\Phi_F)\right] + (\Phi_B^{EVD} - \Phi_F^{EVD})(1-\Phi_F)\right) = 0$$

Breaking down the parenthesis and eliminating common terms we get the following:

$$\left((\Phi_F - \Phi_F^{EVD}) + (\Phi_B^{EVD} - \Phi_B)\right) = 0$$

By changing the notation of the cumulative function, the last expression can be also written as:

$$\left(\Phi_F(\exp(\overline{\theta}_1)) - \Phi_F(\exp(\overline{\theta}_1 - \sigma_F))\right) + \left(\Phi_B(\exp(\overline{\theta}_1 - \sigma_B)) - \Phi_B(\exp(\overline{\theta}_1))\right) = 0$$

The first parenthesis is always positive since $\Phi_F(\exp(\overline{\theta}_1)) > \Phi_F(\exp(\overline{\theta}_1 - \sigma_F))$ because $\exp(\overline{\theta}_1) > \exp(\overline{\theta}_1 - \sigma_F)$. The second parenthesis is always negative since $\Phi_B(\exp(\overline{\theta}_1 - \sigma_B)) < \Phi_B(\exp(\overline{\theta}_1))$ because $\exp(\overline{\theta}_1 - \sigma_B) > \exp(\overline{\theta}_1)$.

Furthermore, since $\sigma_B > \sigma_F$ and since the cumulative function is a monotonic function, the second difference in parenthesis is always larger than the first one, and thus, the whole expression is always negative, as it is shown in Figure 9.

Proof of Proposition 6:

Deriving implicitly equation (22) with respect to W_0 we get the following:

(A.3. 15)

$$\frac{\partial k_1^*}{\partial W_0}=\frac{\frac{\partial\overline{\Theta}_1}{\partial W_0}\left(\overline{\Theta}_1-\frac{(\tilde{\Theta}_{1F}\mid S)}{(1-\Phi_F)}\right)\left(\phi_B-\phi_F\,\frac{(1-\Phi_B)}{(1-\Phi_F)}\right)\beta k_1^{\beta-1}}{-\frac{\partial V(k_1^*,\overline{\Theta}_1)}{\partial^2 k_1^*}}$$

To get to know the sign of this derivative we first need to remember that as it was established in the proof of Proposition 1, for the expected value of the firm to be positive it is necessary that $(\tilde{\Theta}_{1B}\mid S)\geq(1-\Phi_B)\overline{\Theta}_1$. This implies that $\overline{\Theta}_1-\frac{(\tilde{\Theta}_{1F}\mid S)}{(1-\Phi_F)}<0$, and therefore, the first term in parenthesis in the numerator of $\frac{\partial k_1^*}{\partial W_0}$ is always negative.

Since the denominator is positive, ($-\frac{\partial V(k_1^*,\overline{\Theta}_1)}{\partial^2 k_1^*}>0$), and given that $\frac{\partial\overline{\Theta}_1}{\partial W_0}\left(\overline{\Theta}_1-\frac{(\tilde{\Theta}_{1F}\mid S)}{(1-\Phi_F)}\right)>0$ and that $\beta k_1^{\beta-1}>0$, we have that the relationship between the optimal level of capital and internal funds is positive whenever the second parenthesis in the numerator is positive.

In other words, if $\left(\phi_B-\phi_F\frac{(1-\Phi_B)}{(1-\Phi_F)}\right)>0$ then $\frac{\partial k_1^*}{\partial W_0}>0$, and if $\left(\phi_B-\phi_F\,\frac{(1-\Phi_B)}{(1-\Phi_F)}\right)<0$ then $\frac{\partial k_1^*}{\partial W_0}<0$.

We have already seen that whether $(1-\Phi_B)$ is higher or lower than $(1-\Phi_F)$ depends on the value of the given default point, and therefore, on the value of W_0. Something similar happens if we want to know whether ϕ_B is higher or lower than ϕ_F when σ_B is higher than σ_F.

If we derive the density function ϕ_B with respect to σ_B we get the following:

(A.3. 16)

$$\frac{\partial\phi_B}{\partial\sigma_B}=\frac{\phi_B}{\sigma_B}\left(\frac{(\overline{\theta}_1-\theta_B)^2}{\sigma_B^2}-\frac{\sigma_B^2+4}{4}\right)$$

This means that there should be actually two points of $\overline{\theta}_1$ where ϕ_B and ϕ_F intersect each other when the variance of the bank is higher than the variance of the firm.

Therefore, to know whether the function $\left(\phi_B-\phi_F\,\frac{(1-\Phi_B)}{(1-\Phi_F)}\right)$ is positive or negative it is easier to graph it for different parameters of a log-normal distribution, which is what we can see in Figure A. 1. In this figure we can see the values that the function $\left(\phi_B-\phi_F\,\frac{(1-\Phi_B)}{(1-\Phi_F)}\right)$ takes for different log-normal distributions with the same mean but different variance.

In Figure A. 1 we can see that the function $\left(\phi_B - \phi_F \frac{(1-\Phi_B)}{(1-\Phi_F)}\right)$ is always positive for the smallest (negative) values of $\bar{\theta}_1$ and is more likely to be negative at high (positive) values of $\bar{\theta}_1$.

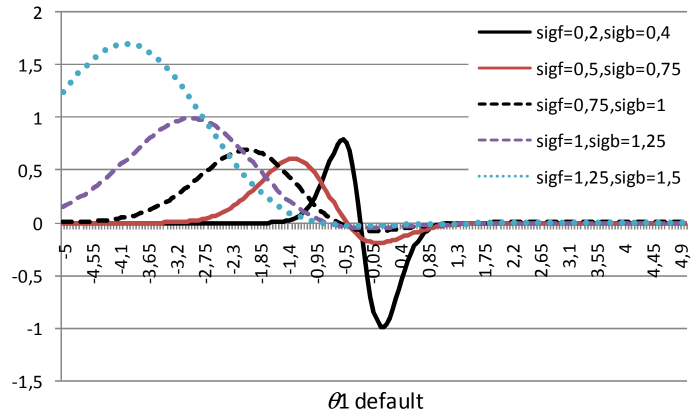


Figure A. 1

What this implies is again that when W_0 is low and therefore $\bar{\theta}_1$ is high, the relationship between investment and wealth should be negative, and when W_0 is high and therefore $\bar{\theta}_1$ is low, the relationship between investment and wealth should be positive.

Proof of Proposition 7

Deriving implicitly the optimal level of capital with respect to internal funds we get the following:

(A.3. 17)

$$\frac{\partial k_1^*}{\partial W_0} = -\frac{\dfrac{\partial V(k_1^*,\overline{\Theta}_1)}{\partial W_0 \partial k_1}}{\dfrac{\partial V(k_1^*,\overline{\Theta}_1)}{\partial^2 k_1^*}} = \frac{-\tau q \dfrac{\partial \Phi_B}{\partial \overline{\Theta}_1}\dfrac{\partial \overline{\Theta}_1}{\partial W_0} - \dfrac{\partial \phi_B}{\partial \overline{\Theta}_1}\dfrac{\partial \overline{\Theta}_1}{\partial W_0}\tau q k_1^* \dfrac{\partial \overline{\Theta}_1}{\partial k_1^*} - \phi_B \tau q k_1^* \dfrac{\partial \overline{\Theta}_1^2}{\partial k_1^* \partial W_0}}{-\dfrac{\partial V(k_1^*,\overline{\Theta}_1)}{\partial^2 k_1^*}}$$

We can remember that $\frac{\partial \overline{\Theta}_1}{\partial k_1^*} = k_1^{-\beta}(1-\Phi_B)^{-1}\left[R - q(1-\tau) - \beta k_1^{\beta-1}\left(\overline{\Theta}_1(1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)\right)\right]$.

Deriving $\frac{\partial \overline{\Theta}_1}{\partial k_1^*}$ with respect to W_0, we get the following:

(A.3. 18)

$$\frac{\partial \overline{\Theta}_1^2}{\partial k_1^* \partial W_0} = k_1^{-\beta}(1-\Phi_B)^{-2}\phi_B \frac{\partial \overline{\Theta}_1}{\partial W_0}\left[R - q(1-\tau) - \beta k_1^{\beta-1}\left(\overline{\Theta}_1(1-\Phi_B) + (\tilde{\Theta}_{1B} \mid D)\right)\right] - k_1^{-\beta}(1-\Phi_B)^{-1}\beta k_1^{\beta-1}(1-\Phi_B)\frac{\partial \overline{\Theta}_1}{\partial W_0}$$

(In the last term I used the fact that $\frac{\partial\left(\overline{\Theta}_1(1-\Phi_B)+(\tilde{\Theta}_{1B}\mid D)\right)}{\partial\overline{\Theta}_1}=(1-\Phi_B)$, according to the proof of Proposition 1 in Appendix A. 3). Simplifying (A.3. 18) we get:

(A.3. 19)
$$\frac{\partial\overline{\Theta}_1^2}{\partial k_1^*\partial W_0}=\frac{\partial\overline{\Theta}_1}{\partial k_1}(1-\Phi_B)^{-1}\phi_B\frac{\partial\overline{\Theta}_1}{\partial W_0}-\beta k_1^{-1}\frac{\partial\overline{\Theta}_1}{\partial W_0}$$

Replacing (A.3. 19) and simplifying again we get:

(A.3. 20)
$$\frac{\partial k_1^*}{\partial W_0}=\frac{-\frac{\partial\overline{\Theta}_1}{\partial W_0}\tau q\left[\phi_B(1-\beta)+\frac{\partial\phi_B}{\partial\overline{\Theta}_1}\frac{\partial\overline{\Theta}_1}{\partial k_1^*}k_1^*+\phi_B^2(1-\Phi_B)^{-1}\frac{\partial\overline{\Theta}_1}{\partial k_1^*}k_1^*\right]}{-\frac{\partial V(k_1^*,\overline{\Theta}_1)}{\partial^2 k_1^*}}$$

Since the denominator is positive, ($-\frac{\partial V(k_1^*,\overline{\Theta}_1)}{\partial^2 k_1^*}>0$), and given that $-\frac{\partial\overline{\Theta}_1}{\partial W_0}>0$, we have that the relationship between the optimal level of capital and internal funds will be positive whenever the term in brackets in the numerator is positive. Since $(1-\beta)>0$ and $\frac{\partial\overline{\Theta}_1}{\partial k_1^*}>0$, the term in brackets will be strictly positive whenever $\frac{\partial\phi_B}{\partial\overline{\Theta}_1}$ is positive. This derivative is equal to:

(A.3. 21)
$$\frac{\partial\phi_B}{\partial\overline{\Theta}_1}=\phi_B\left(\frac{-\overline{\theta}_1+\theta_B-3\sigma_B^2/2}{\overline{\Theta}_1\sigma_B^2}\right)$$

This derivative has a different sign depending on the value of the default point $\overline{\theta}_1$ (or $\overline{\Theta}_1$), and therefore, depends on the value of W_0.

Replacing (A.3. 21) into equation (A.3. 20) we get:

(A.3. 22)
$$\frac{\partial k_1^*}{\partial W_0}=\frac{-\frac{\partial\overline{\Theta}_1}{\partial W_0}\tau q\phi_B\left[(1-\beta)+\frac{\partial\overline{\Theta}_1}{\partial k_1^*}k_1^*\left[\frac{(-\overline{\theta}_1-3\sigma_B^2/2+\theta_B)(1-\Phi_B)+\phi_B\overline{\Theta}_1\sigma_B^2}{\overline{\Theta}_1\sigma_B^2(1-\Phi_B)}\right]\right]}{-\frac{\partial V(k_1^*,\overline{\Theta}_1)}{\partial^2 k_1^*}}$$

Finally, the term in brackets will be positive whenever the following function is positive:

(A.3. 23)
$$(-\overline{\theta}_1-3\sigma_B^2/2+\theta_B)(1-\Phi_B)+\phi_B\overline{\Theta}_1\sigma_B^2$$

In other words, if (A.3. 23) > 0 then $\frac{\partial k_1^*}{\partial W_0}>0$.

Since the second term in the expression is always positive, we have that it is guaranteed that there exists a positive relationship between investment and internal funds whenever $\bar{\theta}_1 < \theta_B - 3\sigma_B^2/2$. This means that at the lowest levels of riskiness (the lowest negative values of $\bar{\theta}_1$) there always exists a positive relationship between investment and internal funds. Since $\bar{\theta}_1$ is smaller the higher is W_0, this means that when W_0 is relatively high, then the relationship between investment and wealth should be positive, i.e. at the highest levels of W_0 we have that $\frac{\partial k_1^*}{\partial W_0} > 0$.

The converse is not true, and it could happen that if $(-\bar{\theta}_1 + \theta_B - 3\sigma_B^2/2)(1-\Phi_B) + \phi_B \bar{\Theta}_1 \sigma_B^2 < 0$ then the relationship between investment capital and wealth may be negative for some range of values of $\bar{\theta}_1$ (or low values of W_0).

Since there is no explicit solution for the default point above which this expression is negative, we can only graph it for different parameters of a log-normal distribution and see when the expression is more likely to be negative. In the following Figure A. 2, I show the graph of $(-\bar{\theta}_1 - 3\sigma_B^2/2 + \theta_B)(1-\Phi_B) + \phi_B \bar{\Theta}_1 \sigma_B^2$ for different distribution functions with different values of σ_B but with the same mean of the distribution $\theta_B = 0$. Therefore, the x-axis represents the default point $\bar{\theta}_1$ which is normally distributed with mean $\theta_B = 0$ and with the respective values of σ_B:

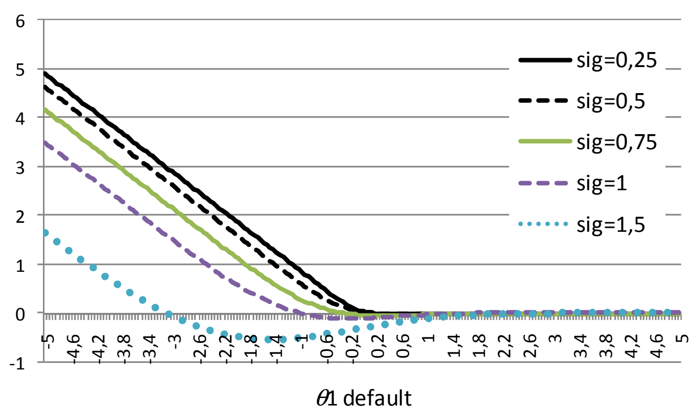


Figure A. 2

We can see in Figure A. 2 that the expression is always positive for low (negative) values of $\bar{\theta}_1$. Therefore, as we said before, when the default point is low (low probability of default) we have that the relationship between investment and internal funds is always positive.

It can be seen in the graph that the expression tends to be negative at the highest levels of $\bar{\theta}_1$, and that the higher the variance, the lowest the point at which the expression can be negative.

This is telling us that the more risky is investment (higher $\overline{\theta}_1$), and the more volatile or uncertain (higher σ_B^2), the more likely that the relationship between investment and internal funds is negative.

Proof of Proposition 8

To prove this proposition we need to derive the probability of default Φ_B with respect to the variance of the bank σ_B: $\frac{\partial \Phi_B}{\partial \sigma_B} = \left(\frac{\partial \Phi_B}{\partial \sigma_B} + \phi_B \frac{\partial \overline{\Theta}_1}{\partial \sigma_B} \right)$ and to use Proposition 4 that says that $\frac{d\overline{\Theta}_1}{d\sigma_B} > 0$.

$$\frac{\partial \Phi_B}{\partial \sigma_B} = \left(\frac{\phi_B}{\sigma_B}\left(-\overline{\theta}_1 + \theta_B + \sigma_B^2/2\right) + \phi_B \frac{\phi_B \overline{\Theta}_1^2 \sigma_B}{(1-\Phi_B)} \right)$$

$$\frac{\partial \Phi_B}{\partial \sigma_B} = \left(-\frac{\phi_B \overline{\theta}_1}{\sigma_B} + \frac{\phi_B^2 \overline{\Theta}_1^2 \sigma_B}{(1-\Phi_B)} + \frac{\phi_B\left(\theta_B + \sigma_B^2/2\right)}{\sigma_B} \right)$$

This derivative is always positive. The sign is given by the term in parenthesis. In Figure A. 3 we can see different examples of this function for different distributions with the same mean, but different variances. We can see that the function defined by the term in parenthesis is always positive no matter what the variance is.

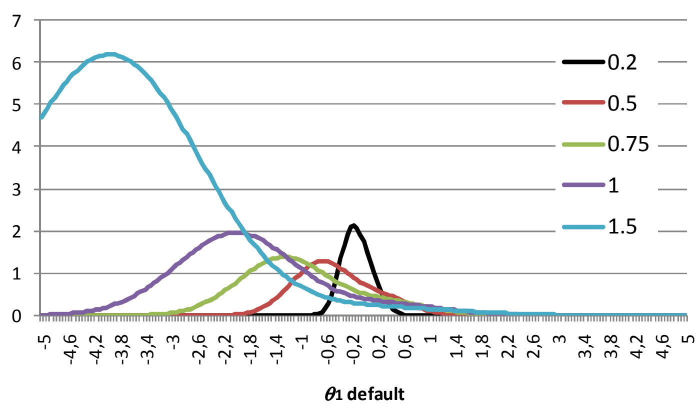


Figure A. 3

Proof of Proposition 9.

The interest rate on a risky loan is given by:

$$R_1 = (1-\Phi_B)^{-1} B_1^{-1} \left[RB_1 - \Phi_B q k_1 (1-\tau) - (\tilde{\Theta}_{1B} \mid D) k_1^{\beta} \right]$$

Deriving with respect to the variance of the bank we get the following:

$$\frac{\partial R_1}{\partial \sigma_B} = (1-\Phi_B)^{-2}\frac{\partial(1-\Phi_B)}{\partial \sigma_B}B_1^{-1}\left[RB_1 - \Phi_B(1-\tau)qk_1 - (\tilde{\Theta}_{1B} \mid D)k_1^{\beta}\right] + ...$$

$$... + (1-\Phi_B)^{-1}B_1^{-1}\left[-\frac{\partial \Phi_B}{\partial \sigma_B}(1-\tau)qk_1 - \frac{\partial(\tilde{\Theta}_{1B} \mid D)}{\partial \sigma_B}k_1^{\beta}\right]$$

Using the results that $\frac{\partial \Phi_B}{\partial \sigma_B} = \frac{\exp(-z_1^2)(\overline{\theta}_1 - \theta_B - \sigma_B^2/2)}{\sqrt{2\pi}\sigma_B^2}$ and that $\frac{\partial(\tilde{\Theta}_{1B} \mid D)}{\partial \sigma_B} = \frac{\overline{\Theta}_1 \exp(-z_1^2)(\overline{\theta}_1 - \theta_B + \sigma_B^2/2)}{\sqrt{2\pi}\sigma_B^2}$ the derivative of the interest rate with respect to the variance is equal to the following:

$$\frac{\partial R_1}{\partial \sigma_B} = -(1-\Phi_B)^{-2}\frac{\exp(-z_1^2)(\overline{\theta}_1 - \theta_B - \sigma_B^2/2)}{\sqrt{2\pi}\sigma_B^2}B_1^{-1}\left[RB_1 - \Phi_B qk_1 - (\tilde{\Theta}_{1B} \mid D)k_1^{\beta}\right] + ...$$

$$... + (1-\Phi_B)^{-1}B_1^{-1}\left[\frac{\exp(-z_1^2)(\overline{\theta}_1 - \theta_B - \sigma_B^2/2)}{\sqrt{2\pi}\sigma_B^2}qk_1 + \frac{\overline{\Theta}_1 \exp(-z_1^2)(\overline{\theta}_1 - \theta_B + \sigma_B^2/2)}{\sqrt{2\pi}\sigma_B^2}k_1^{\beta}\right]$$

Expanding the last expression by separating the term $(\overline{\theta}_1 - \theta_B - \sigma_B^2/2)$ into separated factors $(\overline{\theta}_1 - \theta_B)$ and $(-\sigma_B^2/2)$ and factorizing other common terms, we have the following:

$$= -(1-\Phi_B)^{-1}\frac{\exp(-z_1^2)(\overline{\theta}_1 - \theta_B)}{\sqrt{2\pi}\sigma_B^2}B_1^{-1}\left[\frac{RB_1 - \Phi_B(1-\tau)qk_1 - (\tilde{\Theta}_{1B} \mid D)k_1^{\beta}}{(1-\Phi_B)} - (1-\tau)qk_1 - \overline{\Theta}_1 k_1^{\beta}\right] + ...$$

$$... + (1-\Phi_B)^{-1}\frac{\exp(-z_1^2)(\sigma_B^2/2)}{\sqrt{2\pi}\sigma_B^2}B_1^{-1}\left[\frac{RB_1 - \Phi_B(1-\tau)qk_1 - (\tilde{\Theta}_{1B} \mid D)k_1^{\beta}}{(1-\Phi_B)} - (1-\tau)qk_1 + \overline{\Theta}_1 k_1^{\beta}\right]$$

From equation (17) we can use that $\frac{RB_1 - (1-\tau)qk_1 - (\tilde{\Theta}_{1B} \mid D)k_1^{\beta}}{(1-\Phi_B)} - \overline{\Theta}_1 k_1^{\beta} = 0$, and that $\frac{\Phi_B(1-\tau)qk_1}{(1-\Phi_B)} - (1-\tau)qk_1 = \frac{(1-\tau)qk_1}{(1-\Phi_B)}$ to simplify and get the following expression:

$$\frac{\partial R_1}{\partial \sigma_B} = (1-\Phi_B)^{-1}B_1^{-1}\frac{\exp(-z_1^2)}{\sqrt{2\pi}\sigma_B^2}\frac{\sigma_B^2}{2}\left[\frac{RB_1}{(1-\Phi_B)} - \frac{qk_1}{(1-\Phi_B)} - \frac{(\tilde{\Theta}_{1B} \mid D)}{(1-\Phi_B)}k_1^{\beta} + \overline{\Theta}_1 k_1^{\beta}\right] = (1-\Phi_B)^{-1}B_1^{-1}\frac{\exp(-z_1^2)}{\sqrt{2\pi}\sigma_B^2}\frac{\sigma_B^2}{2}\left[2\overline{\Theta}_1 k_1^{\beta}\right]$$

And therefore we have the result that the risky loans interest rate decreases with the variance and thus with the length of relationship:

$$\frac{\partial R_1}{\partial \sigma_B} = (1-\Phi_B)^{-1}B_1^{-1}\frac{\exp(-z_1^2)}{\sqrt{2\pi}}\left[\overline{\Theta}_1 k_1^{\beta}\right] > 0$$

Appendix A. 4

This appendix shows the algebra needed to get to the first order condition given by equation (20). Maximizing equation (19) with respect to k_1 and rearranging some terms we could find the following first order condition:

(A.4. 1)

$$\left(\frac{\partial(\tilde{\Theta}_{1F}\mid S)}{\partial\overline{\Theta}_1}+\frac{(1-\Phi_F)}{(1-\Phi_B)}\frac{\partial(\tilde{\Theta}_{1B}\mid D)}{\partial\overline{\Theta}_1}\right)\frac{\partial\overline{\Theta}_1}{\partial k_1}k_1^{\beta}+\left((\tilde{\Theta}_{1F}\mid S)+\frac{(1-\Phi_F)}{(1-\Phi_B)}(\tilde{\Theta}_{1B}\mid D)\right)\beta k_1^{\beta-1}=$$

$$+\frac{\partial\overline{\Theta}_1}{\partial k_1}\left(RB_1-qk_1(1-\tau\Phi_B)-(\tilde{\Theta}_{1B}\mid D)k_1^{\beta}\right)\frac{\left[(1-\Phi_F)\phi_B-(1-\Phi_B)\phi_F\right]}{(1-\Phi_B)^2}$$

To simplify the left hand side of this expression I use the following results:

$\frac{\partial(\tilde{\Theta}_{1F}\mid S)}{\partial\overline{\Theta}_1}=-\phi_F\overline{\Theta}_1$ and $\frac{\partial(\tilde{\Theta}_{1B}\mid D)}{\partial\overline{\Theta}_1}=\phi_B\overline{\Theta}_1$

Using those results I can rewrite the left hand side of the first order condition as the following:

$$\left(-\phi_F\overline{\Theta}_1+\frac{(1-\Phi_F)}{(1-\Phi_B)}\phi_B\overline{\Theta}_1\right)\frac{\partial\overline{\Theta}_1}{\partial k_1}.k_1^{\beta}+\left((\tilde{\Theta}_{1F}\mid S)+\frac{(1-\Phi_F)}{(1-\Phi_B)}(\tilde{\Theta}_{1B}\mid D)\right)\beta k_1^{\beta-1}$$

Which after rearranging could be written as follows:

$$\left(\frac{(1-\Phi_F)\phi_B-(1-\Phi_B)\phi_F}{(1-\Phi_B)}\right)\frac{\partial\overline{\Theta}_1}{\partial k_1}\overline{\Theta}_1k_1^{\beta}+\left((\tilde{\Theta}_{1F}\mid S)+\frac{(1-\Phi_F)}{(1-\Phi_B)}(\tilde{\Theta}_{1B}\mid D)\right)\beta k_1^{\beta-1}$$

To simplify the right hand side, I also use that according to the equation for the default point:

$$RB_1-qk_1-(\tilde{\Theta}_{1B}\mid D)k_1^{\beta}=\overline{\Theta}_1k_1^{\beta}(1-\Phi_B)-\tau qk_1$$

Replacing this into the right hand side of the first order condition and rearranging we could get the following right hand side expression:

$$\left((R-q)+\tau q(\Phi_B+\phi_B\frac{\partial\overline{\Theta}_1}{\partial k_1}k_1)\right)\frac{(1-\Phi_F)}{(1-\Phi_B)}+\frac{\partial\overline{\Theta}_1}{\partial k_1}\left(\overline{\Theta}_1k_1^{\beta}-\tau qk_1\right)\frac{\left[(1-\Phi_F)\phi_B-(1-\Phi_B)\phi_F\right]}{(1-\Phi_B)}$$

Cancelling some common terms (the ones that include $\overline{\Theta}_1k_1^{\beta}$) from each side of the equation, and rearranging again we get the following first order condition:

$$\left((\tilde{\Theta}_{1F}\mid S)+\frac{(1-\Phi_F)}{(1-\Phi_B)}(\tilde{\Theta}_{1B}\mid D)\right)\beta k_1^{\beta-1}=$$

$$\left((R-q)+\tau q\Phi_B\right)\frac{(1-\Phi_F)}{(1-\Phi_B)}-\tau qk_1\frac{\partial\overline{\Theta}_1}{\partial k_1}\left(\frac{\left[(1-\Phi_F)\phi_B-(1-\Phi_B)\phi_F\right]}{(1-\Phi_B)}-\frac{(1-\Phi_F)\phi_B}{(1-\Phi_B)}\right)$$

This expression can finally be written as the following:

$$\left((\tilde{\Theta}_{1F}\mid S)+\frac{(1-\Phi_F)}{(1-\Phi_B)}(\tilde{\Theta}_{1B}\mid D)\right)\beta k_1^{\beta-1}=\left((R-q)+\tau q\Phi_B\right)\frac{(1-\Phi_F)}{(1-\Phi_B)}+\tau qk_1\phi_F\frac{\partial\overline{\Theta}_1}{\partial k_1}$$

Which is equal to equation (20).

Alternatively we can derive equation (18) with respect to k_1, and using that $\frac{\partial(\tilde{\Theta}_{1F} \mid S)}{\partial k_1} = -\phi_F \frac{\partial \overline{\Theta}_1}{\partial k_1} \overline{\Theta}_1$ and $\frac{\partial(1-\Phi_F)}{\partial k_1} = -\phi_F \frac{\partial \overline{\Theta}_1}{\partial k_1}$ and equation (A.3. 8), we could get the following FOC:

$$(\tilde{\Theta}_{1F} \mid S)\beta k_1^{\beta-1} + (1-\Phi_F)q - \left(R - q(1-\tau\Phi_B) - (\tilde{\Theta}_{1B} \mid D)\beta k_1^{\beta-1}\right)\frac{(1-\Phi_F)}{(1-\Phi_B)} - \phi_F \frac{\partial \overline{\Theta}_1}{\partial k_1}\left(\overline{\Theta}_1 k_1^{\beta} + qk_1 - R_1 B_1\right) = 0$$

Factorizing common terms and using that according to equation (15) for the default point $\overline{\Theta}_1 k_1^{\beta} + qk_1(1-\tau) - R_1 B_1 = 0$, we get the following FOC, which is equal to equation (20):

$$\left((\tilde{\Theta}_{1F} \mid S) + \frac{(1-\Phi_F)}{(1-\Phi_B)}(\tilde{\Theta}_{1B} \mid D)\right)\beta k_1^{\beta-1} = \left((R-q) + \tau q \Phi_B\right)\frac{(1-\Phi_F)}{(1-\Phi_B)} + \tau q k_1 \phi_F \frac{\partial \overline{\Theta}_1}{\partial k_1}$$

Appendix A. 5

Third case: Asymmetry with bankruptcy costs ($\tau > 0$)

In this final case we can see how asymmetry of information together with bankruptcy costs also generate a non-monotonic relationship between investment and wealth.

In this case the first order condition for optimal investment is given by:

(A.5.1)

$$F(k_1^*, \overline{\Theta}_1) = \left((\tilde{\Theta}_{1F} \mid S) + \frac{(1-\Phi_F)}{(1-\Phi_B)} (\tilde{\Theta}_{1B} \mid D) \right) \beta k_1^{\beta-1} = \left((R-q) + \tau q \Phi_B \right) \frac{(1-\Phi_F)}{(1-\Phi_B)} + \phi_F \frac{\partial \overline{\Theta}_1}{\partial k_1} \tau q k_1 = 0$$

Rearranging and deriving it implicitly with respect to W_0 we get the following (omitting several steps):

(A.5. 2)

$$\frac{\partial k_1^*}{\partial W_0} = \frac{\dfrac{\partial \overline{\Theta}_1}{\partial W_0} [A]}{-\dfrac{\partial F(k_1^*, \overline{\Theta}_1)}{\partial k_1^*}}$$

Where

(A.5. 3)

$$A = \left(\overline{\Theta}_1 - \frac{(\tilde{\Theta}_{1F} \mid S)}{(1-\Phi_F)} \right) \left(\phi_B - \phi_F \frac{(1-\Phi_B)}{(1-\Phi_F)} \right) \beta k_1^{\beta-1} - \tau q \left(\phi_B - \phi_F \beta \frac{(1-\Phi_B)}{(1-\Phi_F)} \right)$$
$$- \tau q k_1^* \phi_F \frac{\partial \overline{\Theta}_1}{\partial k_1^*} \left[\frac{(-\overline{\theta}_1 - 3\sigma_B^2/2 + \tilde{\theta}_{1B})(1-\Phi_B) + \phi_B \overline{\Theta}_1 \sigma_B^2}{\overline{\Theta}_1 \sigma_B^2 (1-\Phi_B)} \right]$$

In this case, the sign of the relationship between investment and wealth depends on the same functions of $\overline{\theta}_1$ defined before. Again, the relationship is positive whenever $\left(\phi_B - \phi_F \frac{(1-\Phi_B)}{(1-\Phi_F)} \right)$ and $\frac{(-\overline{\theta}_1 - 3\sigma_B^2/2 + \tilde{\theta}_{1B})(1-\Phi_B) + \phi_B \overline{\Theta}_1 \sigma_B^2}{\overline{\Theta}_1 \sigma_B^2 (1-\Phi_B)}$ are positive, and the relationship may be negative when these functions are negative.

Appendix A. 6

CHAPTER 5 VARIABLE DEFINITIONS

BANK-FIRM RELATIONSHIP VARIABLES:

AGE: Number of years since the business was [established/purchased/acquired] by the current owner(s).

AGE2: AGE^2.

LN(AGE): Natural logarithm of (1+AGE).

LN(ASSETS): Natural logarithm of the total value of assets.

LENGTHM: Longest length of a relationship with any financial institution (in years) at the time of application.

LENGTH: Length of relationship at the time of application for the most recent loan applied (in months, granted or denied).

LENGTHM/AGE: =(1+LENGTHM)/(1+AGE). It is the length relative to the age of the firm. It may indicate whether there is an asymmetry of information. If LENGTHM=AGE there is symmetry of information (both bank and firm have learned about the firm during the same time). If LENGTHM<AGE, the bank has less information than the firm. As LENGTHM and AGE increase together, it tends to 1.

LENGTHB: =LENGTH if LENGTH< LENGTHM and 0 otherwise.

LN(LENGTHM): Natural logarithm of (1+ LENGTHM).

LN(LENGTH): Natural logarithm of (1+ LENGTH).

HERFINDAHL: Commercial bank deposit herfindahl index of MSA or county where firm's headquarters office is located. Derived from FDIC Summary of deposits data, Dec, 2003. Index equals the sum of the squared market shares times 10,000.

COMPETIT: Dummy variable equal to 1 if Herfindahl Index is smaller than 1000.

INTERM: Dummy variable equal to 1 if Herfindahl Index is between 1000 and 1800.

CONCENTR: Dummy variable equal to 1 if Herfindahl Index is higher than 1800.

CHESAV: Equal to 2 if the firm has a checking and a savings account, equal to 1 if the firm has a checking or a savings account, and equal to 0 if the firm does not have any checking or savings account.

DISTANCE: Approximate number of files from FIRM's main office to the location of institution where firm applied.

NINST: Number of financial institutions the firm has dealing with.

NINSTCHESAV: Number of financial institutions the firm uses for checking and saving services.

NINSTLOAN: Number of financial institutions the firm uses for lending services.

ONEBANK: Dummy variable equal to 1 if the firm has dealings only one financial institution.

BANK: Dummy variable equal to 1 if the most recent loan was from a Bank.

NONFINANC: Dummy variable equal to 1 if the most recent loan was from a non financial institution.

FIRM CHARACTERISCTICS:

PROFITABL: Operating Profits / Equity if Equity>0, or, -Operating Profits / Equity if Equity<0.
INCRSALES: Equal to 1 if Sales have increased during the last 3 years, equal to 0 if Sales have not changed during the last 3 years, and equal to -1 if Sales have decreased during the last 3 years or if firm was not in business three years ago.
LEVERAGE: Total Liabilities / Total Assets.
A0_DB_CREDRK: Dun and Bradstreet Rank Credit Score: - 1 most risky; 6 least risky.
BNKRUPT: Dummy variable equal to 1 if the firm has declared bankruptcy in the last 7 years.
DELINQ: Number of times the firm has been 60 or more days delinquent in their business obligations (including supplier credit).
JUDGMNT: Dummy variable equal to 1 if there are any judgments been rendered against the firm.
AUDITED: Dummy variable equal to 1 if the financial statements or the accounting reports have been audited.
CCORPORATION: Dummy variable equal to 1 if the firm is a C Corporation.
SCORPORATION: Dummy variable equal to 1 if the firm is a S Corporation.
SOLEPROP: Dummy variable equal to 1 if the firm is a firm with a sole proprietor.
PARTNER: Dummy variable equal to 1 if the firm is a Partnership.
URBAN: Dummy variable equal to 1 if the firm is in a MSA and equal to 0 if the firm is in a rural county.
MINING: Dummy variable equal to 1 if SIC code Mining.
CONSTRCT: Dummy variable equal to 1 if SIC code Construction.
MANUFCT: Dummy variable equal to 1 if SIC code Manufacturing.
TRANSP: Dummy variable equal to 1 if SIC code Transportation/Public Utilities.
WHOTRADE: Dummy variable equal to 1 if SIC code Wholesales Trade.
FIRE: Dummy variable equal to 1 if SIC code Finance, Insurance and Real Estate.
SERVICE: Dummy variable equal to 1 if SIC code Service.

<u>OWNER(S) CHARACTERISTICS VARIABLES:</u>

CF_OWNERS: Number of Owners.
ORIGOWN: Dummy variable equal to 1 if the firm was originally established by current owner(s).
EXPER: Weighted average of experience of firm owner(s).
PREVEXPER: Max(EXPER – AGE,0). In other words is the (weighted) previous experience of owners before starting the firm.
OWNAGE: Weighted average age of firm owner(s).
CF_FEMALE: Weighted percentage or ownership that is female.
CF_MINOR: Weighted percentage of ownership that belongs to a minority race.
CF_EDUC: Weighted average education level of owner(s):
1: Less than high school degree (grade 11 or less)
2: High school graduate or equivalent (GED)
3: Some college but no degree granted
4: Associate degree occupational/academic program
5: Trade school/vocational program
6: College degree (BA, BS, AB, etc.)

7: Post graduate degree (MBA,MS,MA,Phd,JD,MD,DDS,etc).

OWNBNKRUPT: Dummy variable equal to 1 if the principal owner has declared bankruptcy in the last 7 years.

OWNDELINQ: Number of times the principal owner has been 60 or more days delinquent in their business obligations (including supplier credit).

OWNJUDGMNT: Dummy variable equal to 1 if there are any judgments been rendered against the principal owner.

OWNHOME: Dummy variable equal to 1 if the principal owner owns a home.

OWNOWORTH: Natural logarithm of principal owner's other assets.

<u>LOAN'S CHARACTERISTICS VARIABLES:</u>

BALANCE: Dummy variable equal to 1 if a compensating balance was required for the most recent line of credit.

GUARANTEE: Dummy variable equal to 1 if the firm was required to have a personal guarantee, cosigner, or other guarantor for the most recent loan/line of credit.

COLLAT: Dummy variable equal to 1 if any type of collateral was required to secure the most
recent loan/line of credit.

COLL_ACCR: Dummy variable equal to 1 if inventory or accounts receivable was required as collateral to secure the most recent loan.

COLL_EQVEH: Dummy variable equal to 1 if vehicles or business equipment was required as collateral to secure the most recent loan.

COLL_SEC: Dummy variable equal to 1 if business securities or deposits was required as collateral to secure the most recent loan.

COLL_BUSREAL: Dummy variable equal to 1 if business real estate was required as collateral to secure the most recent loan.

COLL_PERSREAL: Dummy variable equal to 1 if business real estate was required as collateral to secure the most recent loan.

FIXED: Dummy variable equal to 1 if the interest rate was fixed (not tied to an index).

PRIME01: Dummy variable equal to 1 if the interest rate was tied to the prime rate.

LIBOR01: Dummy variable equal to 1 if the interest rate was tied to the libor rate index.

PRIMERATE: Prime rate at the time of application.

LNMATUR: Log of the maturity term in years if the loan has a fixed maturity.

NEWLOC: Dummy variable equal to 1 if the loan type applied was a new line of credit (granted or denied).

MORTGAGE: Dummy variable equal to 1 if the loan type applied was a mortgage loan (granted or denied).

EQUIPLOAN: Dummy variable equal to 1 if the loan type applied was an equipment loan (granted or denied).

CAPLEASE: Dummy variable equal to 1 if the loan type applied was a capital lease (granted or denied).

RENEWLOC: Dummy variable equal to 1 if the loan type applied was a renewal of a line of credit (granted or denied).

Note: The variables BNKRUPT, DELINQ, JUDGMNT and OWNBNKRUPT, OWNDELINQ and OWNJUDGMNT were collapsed into:

BANKRPT = BNKRUPT + OWNBNKRUPT
DELINQU = DELINQ + OWNDELINQ
JUDGM = JUDGMNT + OWNJUDGMNT

Appendix A. 7

CHAPTER 5 COMPLETE EMPIRICAL RESULTS

Table A.7.1 Complete results from Table 34.

	Whole Sample	**Original Owned**	**Low risk**	**High risk**	**Asymmetric**	**Symmetric**
	Number of obs = 1668	Number of obs = 1040	Number of obs = 626	Number of obs = 414	Number of obs = 531	Number of obs = 435
	F(60, 1607) = 9.25	F(60, 979) = 6.79	F(60, 565) = 7.45	F(60, 353) = 2.83	F(60, 470) = 5.56	F(59, 375) = 4.12
	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000
	R-squared= 0.2614	R-squared = 0.2750	R-squared = 0.3224	R-squared = 0.3525	R-squared = 0.3410	R-squared = 0.3516
	Root MSE = 2.2516	Root MSE = 2.3431	Root MSE = 2.2616	Root MSE = 2.3844	Root MSE = 1.9706	Root MSE = 2.3211
	Whole Sample	**Original Owned**	**Low risk**	**High risk**	**Asymmetric**	**Symmetric**
LNLENGTHM	0.0449	0.3303**	0.4602***	0.1688	0.4018**	-0.6634***
	(0.627)	(0.031)	(0.008)	(0.574)	(0.039)	(0.005)
LNAGE	-0.2934**	-0.6019***	-0.5529**	-0.5881*	-0.6573**	
	(0.011)	(0.001)	(0.021)	(0.078)	(0.020)	
LNAMOUNT	-0.3057***	-0.4093***	-0.4549***	-0.2902**	-0.3674***	-0.5241***
	(0.000)	(0.000)	(0.000)	(0.015)	(0.000)	(0.000)
LNASSETS	-0.1018**	-0.0312	-0.0308	-0.0915	-0.0715	-0.0107
	(0.050)	(0.656)	(0.714)	(0.455)	(0.339)	(0.934)
DISTANCE	0.0008	0.0011	0.0025***	-0.0007	0.0015**	0.0000
	(0.178)	(0.152)	(0.009)	(0.345)	(0.019)	(0.994)
CHESAV	0.0663	-0.1185	-0.1195	-0.1714	0.0000	-0.1714
	(0.574)	(0.434)	(0.533)	(0.526)	(1.000)	(0.501)
NINSTLOANS	0.0469*	0.0127	-0.0389	0.0548	-0.0430	0.1278*
	(0.097)	(0.754)	(0.560)	(0.378)	(0.367)	(0.094)
ONEBANK	-0.1363	-0.4126*	-0.2367	-0.9179**	-0.2777	-0.7779*
	(0.437)	(0.087)	(0.445)	(0.045)	(0.379)	(0.052)
BANK	-0.2915	-0.2556	0.3807	-0.9347*	0.1352	-0.0475
	(0.277)	(0.468)	(0.357)	(0.078)	(0.734)	(0.929)
INTERM	-0.3643	-0.4808	-0.4569	-0.5967	-0.3543	-1.2729*
	(0.140)	(0.180)	(0.388)	(0.165)	(0.304)	(0.060)
CONCENTR	-0.1816	-0.3307	-0.4727	-0.4807	-0.4390	-0.8680
	(0.483)	(0.381)	(0.381)	(0.299)	(0.206)	(0.208)
PROFITABL	-0.0034	-0.0004	-0.0724**	-0.0058	-0.0007	0.0020
	(0.528)	(0.961)	(0.024)	(0.488)	(0.953)	(0.569)
A0_DB_CREDRK	-0.0679	-0.0563	-0.0314	-0.1079	-0.1416**	0.2182**
	(0.154)	(0.341)	(0.694)	(0.249)	(0.045)	(0.014)
CF_OWNERS	0.0006	0.0079	0.0031	0.0553*	-0.0050	0.0252*
	(0.195)	(0.187)	(0.621)	(0.057)	(0.335)	(0.059)
AUDITED	-0.0946	-0.3861	-0.3628	-0.2636	-0.6183*	0.5170
	(0.620)	(0.151)	(0.312)	(0.534)	(0.072)	(0.333)
URBAN	-0.0458	-0.0368	0.0874	-0.4446	-0.1726	0.1633
	(0.762)	(0.854)	(0.719)	(0.276)	(0.484)	(0.592)
BANKRPT	0.5223	0.7481	1.6224**	0.0137	0.9076*	0.3713
	(0.317)	(0.296)	(0.016)	(0.989)	(0.090)	(0.715)
DELINQU	-0.0025	-0.0304	-0.0379	0.0226	0.0863	-0.0413
	(0.956)	(0.587)	(0.659)	(0.782)	(0.286)	(0.592)
JUDGM	0.0621	0.2577	0.4811	0.2617	0.2591	0.5273*
	(0.761)	(0.325)	(0.144)	(0.514)	(0.506)	(0.095)
OWNMANAG	0.0054	-0.0027	-0.0452	0.0940	-0.1982	0.4181
	(0.968)	(0.989)	(0.849)	(0.815)	(0.457)	(0.180)
PREVEXPER	-0.0035	-0.0118	-0.0060	-0.0203	-0.0241	-0.0112
	(0.705)	(0.329)	(0.685)	(0.305)	(0.191)	(0.519)
LNOWNAGE	-0.5781	-0.6269	-0.9473	-0.3927	0.3929	-0.5268
	(0.161)	(0.270)	(0.154)	(0.702)	(0.612)	(0.495)
CF_FEMALE	-0.0039**	-0.0049**	-0.0032	-0.0067	-0.0057**	-0.0038

	Whole Sample	Original Owned	Low risk	High risk	Asymmetric	Symmetric
	(0.036)	(0.041)	(0.239)	(0.134)	(0.038)	(0.313)
CF_MINOR	0.0029	0.0023	0.0028	0.0040	-0.0015	0.0155**
	(0.385)	(0.592)	(0.598)	(0.619)	(0.778)	(0.012)
CF_EDUC	-0.1091***	-0.1180***	-0.0844	-0.1598**	-0.0996*	-0.1281*
	(0.001)	(0.010)	(0.119)	(0.046)	(0.091)	(0.076)
OWNHOME	-0.7924***	-0.5510*	-0.2951	-0.7771	-0.4866	-0.2821
	(0.009)	(0.090)	(0.420)	(0.136)	(0.154)	(0.544)
OWNOWORTH	-0.0353	-0.0215	-0.0568	0.0168	0.0167	-0.0439
	(0.203)	(0.506)	(0.119)	(0.730)	(0.667)	(0.352)
BALANCE	0.0593	0.2478	0.6211	-0.5628	0.7760**	0.0122
	(0.800)	(0.438)	(0.129)	(0.343)	(0.044)	(0.984)
GUARANTEE	-0.0857	0.0414	0.2494	-0.4756*	0.2669	-0.4653*
	(0.481)	(0.795)	(0.196)	(0.089)	(0.204)	(0.061)
COLLAT	-0.3362**	-0.3234*	-0.2716	-0.6253*	-0.2528	-0.2798
	(0.026)	(0.096)	(0.284)	(0.079)	(0.342)	(0.343)
COLL_ACCR	0.3552**	0.2676	0.3526	0.0616	0.2051	0.3358
	(0.030)	(0.220)	(0.214)	(0.872)	(0.466)	(0.391)
COLL_SEC	-0.0126	0.1284	0.2156	-0.0775	0.1360	0.0338
	(0.950)	(0.655)	(0.568)	(0.884)	(0.661)	(0.948)
COLL_BUSREAL	0.6262***	0.3727	0.5067	0.2443	0.1000	0.6931
	(0.000)	(0.142)	(0.119)	(0.609)	(0.766)	(0.118)
COLL_PERSR~L	0.0267	-0.0230	-0.0300	0.1409	0.1689	-0.3360
	(0.878)	(0.918)	(0.921)	(0.722)	(0.599)	(0.387)
COLL_OTHEP~S	-0.2033	-0.4345	-1.2721**	0.4216	-0.3362	-0.4700
	(0.457)	(0.228)	(0.036)	(0.436)	(0.497)	(0.514)
COLL_OTHER	0.2969	0.7274	-0.0868	1.2301	0.5456	2.2280***
	(0.453)	(0.209)	(0.874)	(0.281)	(0.432)	(0.009)
FIXED	1.0564***	0.9341***	0.8544***	0.9541***	0.9210***	0.5559**
	(0.000)	(0.000)	(0.000)	(0.003)	(0.000)	(0.050)
PRIMERATE	0.1054	0.1019	0.0312	0.3698*	0.2618**	-0.1139
	(0.123)	(0.239)	(0.747)	(0.074)	(0.028)	(0.410)
NEWLOC	-0.5188***	-0.4464**	-0.3723	-0.5298	-0.2762	-0.9366***
	(0.001)	(0.022)	(0.118)	(0.118)	(0.244)	(0.005)
CAPLEASE	0.1327	0.3924	-0.2076	0.2598	0.6108	0.1341
	(0.734)	(0.420)	(0.645)	(0.780)	(0.443)	(0.843)
MORTGAGE	-0.0971	0.0260	-0.0813	-0.1941	0.7964**	-0.9581**
	(0.650)	(0.931)	(0.833)	(0.718)	(0.047)	(0.044)
VEHICLE	-1.5389***	-1.4579***	-1.3996***	-1.2551**	-1.0045**	-1.6684***
	(0.000)	(0.000)	(0.002)	(0.034)	(0.032)	(0.000)
EQUIPLOAN	-0.6262***	-0.6367**	-0.8739**	0.1587	-0.1410	-0.9127*
	(0.004)	(0.036)	(0.014)	(0.783)	(0.704)	(0.071)
OTHERTYPE	-0.0319	0.0035	0.5764	-0.7561	-0.4221	0.0711
	(0.929)	(0.994)	(0.364)	(0.202)	(0.380)	(0.904)
LNMATUR	1.8865**	1.8538*	1.0371	2.8947	1.3625	0.0279
	(0.015)	(0.055)	(0.227)	(0.136)	(0.304)	(0.980)
_cons	14.4279***	15.3241***	16.9374***	13.4317***	9.7674***	18.8175***
	(0.000)	(0.000)	(0.000)	(0.003)	(0.002)	(0.000)

***,**,* indicate significance at 1%, 5% and 10% respectively. P-values are in parenthesis. All the regressions include also the control variables MINING, CONSTRCT, MANUFCT, TRANSP, WHOTRADE, FIRE, SERVICE and 8 regional dummies.

Table A.4.2 Complete results from Table 35.

	Asymmetric		Symmetric	
	Number of obs = 334	Number of obs = 233	Number of obs = 197	Number of obs = 178
	F(59, 273) = .	F(58, 174) = 4.08	F(59, 136) = .	F(58, 119) = 3.96
	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000	Prob > F = 0.0000
	R-squared = 0.4145	R-squared = 0.4258	R-squared = 0.4578	R-squared = 0.5284
	Root MSE = 1.9038	Root MSE = 2.23	Root MSE = 2.0635	Root MSE =2.367
	Low Risk	**High Risk**	**Low Risk**	**High Risk**
LNLENGTHM	0.4951**	-0.4585	0.2257	-0.9689**
	(0.025)	(0.239)	(0.586)	(0.012)
LNAGE	-0.6804**		-0.6567	
	(0.023)		(0.293)	
LNAMOUNT	-0.4352***	-0.4557**	-0.1617	-0.4358**
	(0.003)	(0.012)	(0.310)	(0.032)
LNASSETS	-0.0611	-0.1928	-0.2377*	-0.0095
	(0.528)	(0.294)	(0.082)	(0.961)
DISTANCE	0.0019**	0.0000	0.0003	-0.0002
	(0.038)	(0.989)	(0.695)	(0.883)
CHESAV	0.1041	-0.3569	-0.2744	0.1357
	(0.675)	(0.299)	(0.448)	(0.793)
NINSTLOANS	-0.1102	0.0333	0.0332	0.2672*
	(0.163)	(0.769)	(0.714)	(0.070)
ONEBANK	-0.1868	-1.2177**	-1.2567	0.3611
	(0.605)	(0.030)	(0.138)	(0.588)
BANK	0.8456	0.3416	-0.4565	-0.6698
	(0.137)	(0.603)	(0.403)	(0.401)
INTERM	0.0716	-2.3028**	-1.1176*	-0.6874
	(0.883)	(0.048)	(0.081)	(0.368)
CONCENTR	-0.2178	-1.7556	-0.7159	-1.0351
	(0.652)	(0.126)	(0.317)	(0.166)
PROFITABL	-0.0672	-0.0839	-0.0016	0.0024
	(0.149)	(0.303)	(0.914)	(0.568)
A0_DB_CREDRK	-0.0890	0.2378*	-0.1987	0.2553
	(0.328)	(0.063)	(0.148)	(0.120)
CF_OWNERS	-0.0029	0.0124	0.0038	0.1276***
	(0.663)	(0.191)	(0.884)	(0.001)
AUDITED	-0.7034**	0.6270	0.1642	0.3825
	(0.035)	(0.480)	(0.817)	(0.657)
URBAN	-0.1154	0.6008	-0.7794	-0.6549
	(0.659)	(0.117)	(0.230)	(0.373)
BANKRPT	0.4555	1.3297	1.1825	-0.0205
	(0.415)	(0.131)	(0.222)	(0.991)
DELINQU	0.1074	0.0641	0.1469	-0.0086
	(0.385)	(0.644)	(0.285)	(0.946)
JUDGM	0.2345	1.1888**	0.6660	0.2373
	(0.618)	(0.025)	(0.276)	(0.682)
OWNMANAG	-0.2872	0.6707	0.1510	0.4909
	(0.345)	(0.127)	(0.793)	(0.466)
PREVEXPER	-0.0255	-0.0102	-0.0261	-0.0584**
	(0.237)	(0.709)	(0.487)	(0.033)
LNOWNAGE	0.4666	-0.2336	0.3476	0.5526
	(0.559)	(0.844)	(0.817)	(0.709)
CF_FEMALE	-0.0056*	0.0044	-0.0066	-0.0115**
	(0.083)	(0.445)	(0.294)	(0.049)
CF_MINOR	0.0048	0.0064	-0.0102	0.0297**
	(0.481)	(0.406)	(0.190)	(0.015)
CF_EDUC	-0.0720	-0.1033	0.0118	-0.2362*
	(0.328)	(0.267)	(0.919)	(0.075)
OWNHOME	-0.9216**	-0.1854	-0.0689	-0.4219
	(0.042)	(0.737)	(0.904)	(0.653)
OWNOWORTH	-0.0348	-0.0226	0.0854	-0.0090
	(0.541)	(0.779)	(0.163)	(0.910)
BALANCE	1.2836***	-0.4724	-0.3382	0.3637
	(0.004)	(0.657)	(0.701)	(0.731)
GUARANTEE	0.1982	-0.1782	0.1388	-0.9508**
	(0.440)	(0.609)	(0.748)	(0.031)

	Asymmetric		Symmetric	
COLLAT	-0.4959	0.4566	-0.1330	-1.5266**
	(0.155)	(0.243)	(0.786)	(0.015)
COLL_ACCR	0.7281**	0.0365	-0.7414	0.6382
	(0.042)	(0.954)	(0.208)	(0.317)
COLL_SEC	0.1818	-0.0398	0.2226	-0.7620
	(0.673)	(0.955)	(0.716)	(0.441)
COLL_BUSREAL	0.4428	0.2399	-0.0402	0.8770
	(0.295)	(0.667)	(0.954)	(0.308)
COLL_PERSR~L	0.2153	-1.2758**	0.0432	1.3106**
	(0.592)	(0.018)	(0.932)	(0.042)
COLL_OTHEP~S	-0.8341	-1.3270	0.1329	1.6255
	(0.298)	(0.224)	(0.855)	(0.143)
COLL_OTHER	-0.7782	2.0343**	0.3030	
	(0.241)	(0.042)	(0.809)	
FIXED	0.8189***	0.5721	1.3875***	0.7422
	(0.005)	(0.156)	(0.002)	(0.153)
PRIMERATE	0.2327*	-0.2053	0.5073*	0.2494
	(0.100)	(0.179)	(0.074)	(0.439)
NEWLOC	-0.1427	-0.8595	-0.2227	-0.4329
	(0.614)	(0.109)	(0.602)	(0.441)
CAPLEASE	-0.2733	-0.8810	1.2458	-0.6638
	(0.673)	(0.495)	(0.437)	(0.524)
MORTGAGE	0.6417	-0.9130	0.8511	-1.8461*
	(0.221)	(0.210)	(0.225)	(0.065)
VEHICLE	-0.5444	-2.1573***	-1.5369*	-0.7740
	(0.374)	(0.003)	(0.055)	(0.401)
EQUIPLOAN	-0.1207	-1.3879**	-0.1774	0.5402
	(0.802)	(0.018)	(0.807)	(0.623)
OTHERTYPE	0.0803	0.8777	-1.7437*	-0.8685
	(0.879)	(0.311)	(0.095)	(0.308)
LNMATUR	1.8201	0.7930	2.6225	-0.3560
	(0.112)	(0.546)	(0.373)	(0.850)
_cons	9.5299***	17.8213***	8.0054	12.9768**
	(0.003)	(0.000)	(0.180)	(0.043)

***,**,* indicate significance at 1%, 5% and 10% respectively. P-values are in parenthesis. All the regressions include also the control variables MINING, CONSTRCT, MANUFCT, TRANSP, WHOTRADE, FIRE, SERVICE and 8 regional dummies.

REFERENCES

- Abel. A.B. and Blanchard, O.J. (1986), "The Present Value of Profits and Cyclical Movements in Investment" Econometrica, Vol. 54, No. 2 (Mar., 1986), pp. 249-273.
- Aghion, P. and Bolton, P., (1992) "An Incomplete Contracts Approach to Financial Contracting", The Review of Economic Studies, Vol. 59, No. 3 (Jul., 1992), pp. 473-494.
- Albuquerque, R. and Hopenhayn, H. A. (2004), Optimal Lending Contracts and Firm Dynamics. Review of Economic Studies, 71: 285–315.
- Allayannis, George and Mozumdar, Abon. "The Impact of Negative Cash Flow and Influential Observations on Investment–Cash Flow Sensitivity Estimates." Journal of Banking and Finance, 2004, 28(5), pp. 901–30.
- Almeida, H. and Campello, M. (2005), "External Financing and Financing Constraints", mimeo, New York University.
- Almeida, H. and Campello, M., (2002), Financial Constraints and Investment Cash-Flow Sensitivities: New Research Directions", mimeo, New York University.
- Almeida, H. and Campello, M., (2007), "Financial Constraints, Asset Tangibility, and Corporate Investment", Review of Financial Studies, Volume20, Issue5, Pp. 1429-1460.
- Alti, A. (2003), "How Sensitive Is Investment to Cash Flow When Financing Is Frictionless?, The Journal of Finance, Volume 58, Issue 2, pages 707–722, April 2003.
- Andrade G. and Kaplan S. N., (1998) "How Costly is Financial (not Economic) Distress? Evidence from Highly Leveraged Transactions that Became Distressed," Journal of Finance 53 (October 1998), pp. 1443–1493.
- Angelini, P., Di Salvo, R. and Ferri, G. (1998) "Availability and cost of credit for small businesses: Customer relationships and credit cooperatives", Journal of Banking & Finance, Volume 22, Issues 6-8, August 1998, Pages 925-954.
- Beck, T., Demirgüç-Kunt, A., Laeven, L. and Maksimovic V., (2006) "The determinants of financing obstacles" Journal of International Money and Finance 25 (2006), 932-952.
- Beck, T., Demirgüç-Kunt, A., Maksimovic V., (2002) "Financial and Legal Constraints to Firm Growth, Does Size Matter?" World Bank, 2002
- Beck, T., Demirgüç-Kunt, A., Maksimovic V., (2004) " Bank Competition and Access to Finance: International Evidence", Journal of Money, Credit and Banking, Vol. 36, No. 3, June 2004.
- Beck, T., Levine, R. And Loayza, N. (2000) "Finance and the Sources of Growth" Journal of Financial Economics, 58(1).
- Berger, A. N. and Udell, G. (1995), "Relationship lending and lines of credit in Small Firm Finance", Journal of Business 3(68), 351-381.
- Berger, A. N., G. F. Udell (1995), "Relationship Lending and Lines of Credit in Small Firm Finance", Journal of Business, 1995, vol 68, no. 3.
- Berger, A. N., G. F. Udell (1998), "The Economics of Small Business Finance: The Roles of Private Equity and Debt Markets in the Financial Growth Cycle", Journal of Banking and Finance, Vol. 22, 1998.

- Berger, A. N., G. F. Udell (2000), “Small Business and Debt Finance”, Handbook of Entrepreneurship, November 2000.
- Berger, A. N., G. F. Udell (2002), “Small Business Credit Availability and Relationship Lending: The Importance of Bank Organizational Structure”, Economic Journal, 2002.
- Berkowitz, J., White, M., “Bankruptcy and Small Firms’ Access to Credit”, University of Michigan 2000.
- Berlin, M. and Mester, L.J. (1988), “On the profitability and cost of relationship lending”, Journal of Banking & Finance, Volume 22, Issues 6-8, August 1998, Pages 873-897.
- Bernanke, B., Gertler, M. (1989). “Agency Costs, Net Worth, and Business Fluctuations”. American Economic Review 79:14-31.
- Bernanke, B., Gertler, M. (1990). “Financial Fragility and Economic Performance”. Quarterly Journal of Economics 105:87-114.
- Bernardo, J. M. , Smith, A. F. (1994), “Bayesian Theory” Wiley, 1994.
- Besanko, D. and Thakor, A.V. (1986). “Collateral and Rationing: Sorting Equilibria in Monopolistic and Competitive Credit Markets”, International Economic Review 28:671-89.
- Bester, H. (1985). “Screening vs. Rationing in Credit Markets with Imperfect Information” American Economic Review 75:850-855, 1985.
- Black, S., Strahan, P. (2002), “Entrepreneurship and Bank Credit Availability”, The Journal of Finance, Vol LVII, No. 6, December 2002.
- Blackwell, D.W., and Winters, DB. (1997), “Banking relationships and the effect of monitoring on loan pricing”, Journal of Financial Research, 1997.
- Blair, M. and Wallman, S. (2001), “Unseen Wealth” Published by Brookings Institution.
- Blanchflower, D., Levine, P. and Zimmerman, D. (1998), “Discrimination in the Small-Business Credit Market”, The Review of Economics and Statistics, November 2003, Vol. 85, No. 4, Pages 930-943.
- Bolton, P. and Scharfstein, D.S. (1990), “A Theory of Predation Based on Agency Problems in Financial Contracting”. The American Economic Review, Vol. 80, No. 1 (Mar., 1990), pp. 93-106.
- Bond, S. and Cummins, J., “Noisy share prices and the Q model of investment”, Institute for Fiscal Studies, Discussion Paper, 2001
- Bond, S., and C. Meghir (1994), “Dynamic investment models and the firm’s financial policy”, Review of Economic Studies 61:197–222.
- Boot, A. W. A. (2000), “Relationship banking: What do we know?”, Journal of Financial Intermediation 9, 7-25.
- Boot, A. W. A. and Thakor, A. V. (1994), “Moral hazard and secured lending in an infinitely repeated credit market game”, International Economic Review 35(4), 899-920.
- Boot, A. W. A. and Thakor, A. V. (2000), “Can relationship banking survive competition?”, Journal of Finance 55(2), 679-713.
- Boyd, J.H. and Prescott, E. (1986), “Financial intermediary-coalitions”, Journal of Economic Theory, Volume 38, Issue 2, April 1986, Pages 211-232.

- Brealey, R. and Myers, S. (2003), “Principles of Corporate Finance”, Seventh Edition, The MacGrow-Hill Companies 2003.
- Caggese, A. (2007) “Testing financing constraints on firm investment using variable capital” Journal of Financial Economics Volume 86, Issue 3, December 2007, Pages 683-723.
- Caggese, A., (2005) “Financing Imperfections and the Investment Decisions of Small Firms”. Working Paper, Pompeu Fabra University.
- Calomiris, C. And Hubbard, G. (1990). “Firm Heterogeneity, Internal Finance, and Credit Rationing” Economic Journal 100:90-104.
- Campello, M., Graham, JR. and Harvey, CR. (2010), “The real effects of financial constraints: Evidence from a financial crisis”, Journal of Financial Economics 2010.
- Cavalluzzo, K. Cavalluzzo, LC and Wolken JD (2002), “Competition, Small Business Financing, and Discrimination: Evidence from a New Survey”, The Journal of Business, Vol. 75, No. 4 (Oct., 2002), pp. 641-679.
- Cetorelli, N. (2004), “Real effects of bank competition” Journal of Money, Credit & Banking, Vol. 36, 2004.
- Cetorelli, N. and Gambera, M. (2001). "Banking Market Structure, Financial Dependence and Growth: International Evidence from Industry Data." Journal of Finance 56, 617-648.
- Cetorelli, N. and Peretto, PF. (2000). "Oligopoly Banking and Capital Accumulation." Federal Reserve Bank of Chicago Working Paper No. 2000-12.
- Cleary, S. (1999), “The Relationship between Firm Investment and Financial Status”, Journal of Finance, Vol. LIV, No.2, April 1999.
- Cleary, S., Povel, P. and Raith, M. (2009), “The U-Shaped Investment Curve: Theory and Evidence” Journal of Financial and Quantitative Analysis (2007), 42: 1-39 Cambridge University Press.
- Clementi, G.L. and Hopenhayn, H.A. "A Theory of Financing Constraints and Firm Dynamics," GSIA Working Papers 2002-E9, Carnegie Mellon University.
- Cooley, T. and Quadrini, V., (2001) “Financial Markets and Firm Dynamics” The American Economic Review, Vol. 91, December 2001, pp. 1286-1310.
- Cooper, R. and Ejarque, J. (2003), “Financial frictions and investment: requiem in Q”, Review of Economic Dynamics, Volume 6, Issue 4, October 2003, Pages 710-728.
- Cressy, Robert, (1996), “Are Business Startups Debt-Rationed?”, The Economic Journal, Vol 106, No. 438, Sep 1996, 1253-1270.
- Cummins, J.G., Hassett, K.A. and Oliner, S.D., (2006), “Investment Behavior, Observable Expectations and Internal Funds” The American Economic Review, Volume 96, Number 3, June 2006 , pp. 796-810(15).
- De Meza, D., Webb, D. (1987), “Too Much Investment: A Problem of Asymmetric Information”, The Quarterly Journal of Economics, Vol 102, No. 2 (May, 1987), 281-292.
- De Meza, Southey (1996), “The Borrower’s Curse: Optimism, Finance and Entrepreneurship” The Economic Journal, Vol 106, No. 435, Mar 1996, 375-386
- Degryse, H. and Van Cayseele, P. (2000), “Relationship lending within a bank-based system: Evidence from european small business data”, Journal of Financial Intermediation 9(1), 90-109.

- Dell'Ariccia, G. (2001), "Asymmetric information and the structure of the banking industry", European Economic Review 45(10), 1957-1980.
- Demirguc-Kunt, A., and V. Maksimovic (1998), "Law, finance, and firm growth", Journal of Finance, 53:2107–2137.
- Demirgüç-Kunt, A., Laeven, L. and Levine, R. (2003), "Regulations, Market Structure, Institutions, and the Cost of Financial Intermediation" NBER, Working Paper No. W9890.
- Devereux, M., Schiantarelli, F., 1990. Investment, financial factors and cash flow from UK panel data. In: Hubbard, G.(Ed.), Information, Capital Markets and Investment. University of Chicago Press, Chicago, IL.
- DeYoung, R., Goldberg, LG. and White, LJ, (1999), "Youth, adolescence, and maturity of banks: Credit availability to small business in an era of banking consolidation", Journal of Banking & Finance, Volume 23, Issues 2-4, February 1999, Pages 463-492.
- Diamond, D. (1991), "Debt maturity structure and liquidity risk", Quarterly Journal of Economics 106:709–737.
- Diamond, D.W. (1984), "Financial Intermediation and Delegated Monitoring", The Review of Economic Studies, Vol. 51, No. 3 (Jul., 1984), pp. 393-414.
- Dinç, Serdar I. (2000). "Bank Reputation, Bank Commitment, and the Effects of Competition in Credit Markets." Review of Financial Studies 13, 781-812.
- Elsas R. and Krahnen JP., (1998) "Is Relationship Lending Special? Evidence from Credit-File Data in Germany", Journal of Banking & Finance Volume 22, Issues 10-11, October 1998, Pages 1283-1316.
- Elsas, R. (2005), "Empirical determinants of relationship lending", Journal of Financial Intermediation 14, 32-57.
- Elyasiani, E. and Goldberg, LG (2004), "Relationship lending: a survey of the literature", Journal of Economics and Business, Volume 56, Issue 4, July-August 2004, Pages 315-330.
- Erickson, T., and T. Whited (2000), "Measurement error and the relationship between investment and q", Journal of Political Economy 108:1027–1057.
- Evans, D. (1987), "The Relationship between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries". The Journal of Industrial Economics, Vol. 35, No. 4, The Empirical Renaissance in Industrial Economics, (Jun., 1987), pp. 567-581.
- Fama, E. (1985), "What's different about banks?", Journal of Monetary Economics 15, 29-36.
- Fazzari, S., Hubbard, G., and Petersen, B. (1988). "Financing Constraints and Corporate Investment" NBER Working Paper No. 2387.
- Ferri, G. and Messori, M. (2000), "Bank–firm relationships and allocative efficiency in Northeastern and Central Italy and in the South", Journal of Banking & Finance, Volume 24, Issue 6, June 2000, Pages 1067-1095
- Freixas, X. (2005), `Deconstructing relationship banking', Investigaciones Economicas 29(1), 3-31.
- Gale, D. and Hellwig, M. (1985), "Incentive-Compatible Debt Contracts: The One-Period Problem". Review of Economic Studies 1985, LII, 647-663.
- Galindo, A., Schiantarellli, F. (2002), "Credit Constraints in Latin America: An Overview of the Micro Evidence". Research Network Working Paper #R-472, Inter-American Development Bank, September 2002.

- Gehrig, T. (1998), "Screening, cross-border banking, and the allocation of credit", Research in Economics, Volume 52, Issue 4, December 1998, Pages 387-407.
- Gelos, G., Werner, A., 2002. Financial liberalization, credit constraints, and collateral: investment in the Mexican manufacturing sector. Journal of Development Economics 67 (1), 1-27.
- Gertler, M., (1992). "Financial Capacity and Output Fluctuations in an Economy with Multi-Period Financial Relationships." Review of Economic Studies 59:455-72, 1992.
- Gilchrist, S., and C.P. Himmelberg (1995), "Evidence on the role of cash flow for investment", Journal of Monetary Economics 36:531–572.
- Gomes, J.F. (2001), "Financing Investment." American Economic Review, 2001, 91(5), pp. 1263–85.
- Greenacre, M.J. (1993) "Correspondence Analysis in Practice". Academic Press, London.
- Greenbaum, S. I., Kanatas, G. and Venezia, I. (1989), "Equilibrium loan pricing under the bank-client relationship", Journal of Banking and Finance 13(2), 221-235.
- Greenwald, B., J.E. Stiglitz and A. Weiss (1984), "Informational imperfections in the capital market and macroeconomic fluctuations", American Economic Review, 74:194–199.
- Guzman, Mark G. (2000). "Bank Structure, Capital Accumulation and Growth: A Simple Macroeconomic Model." Economic Theory 16, 421-455.
- Hall, B. (1987) "The Relationship Between Firm Size and Firm Growth in the US Manufacturing Sector", The Journal of Industrial Economics, Vol. 35, No. 4, The Empirical Renaissance in Industrial Economics, (Jun., 1987), pp. 583-606.
- Hannan. T. and Berger, AN. (1991), "The Rigidity of Prices: Evidence from the Banking Industry", The American Economic Review, Vol. 81, No. 4 (Sep., 1991), pp. 938-945.
- Harhoff, D. and Körting, T. (1998), "Lending relationships in Germany – Empirical evidence from survey data", Journal of Banking & Finance Volume 22, Issues 10-11, October 1998, Pages 1317-1353.
- Hart, O., and J. Moore (1994), "A theory of debt based on the inalienability of human capital", Quarterly Journal of Economics 109:841–879.
- Hart, O., and J. Moore (1998), "Default and renegotiation: a dynamic model of debt", Quarterly Journal of Economics 113:1–41.
- Hayashi, F. (1982), "Tobin's marginal q and average q: A neoclassical interpretation", Econometrica, Vol. 50, No. 1 (Jan., 1982), pp. 213-22.
- Heider, F. (2002), "Rationing despite Screening: A motif-based on price discrimination" mimeo, Stern School of Business, New York University.
- Heider, F. and Halov, N. (2005), "Capital structure, risk and asymmetric information"
- Hennessy, C.A., Levy, A. and Whited, T.M. (2005), "Testing Q Theory with Financing Frictions" , mimeo, University of California at Berkeley.
- Hoshi, T., Kashyap, A., Scharfstein, D., 1991. Corporate structure, liquidity and investment: evidence from Japanese Industrial Groups. Quarterly Journal of Economics 106 (1), 33-60.
- Houston, J. F. and James, C. M. (2001), "Do relationships have limits? Banking relationships, financial constraints, and investment", Journal of Business 74(3), 347-74.

- Hovakimian, G. (2009), "Determinants of Investment Cash Flow Sensitivity", Financial Management, Volume 38, Issue 1, pages 161–183, Spring 2009.
- Hubbard, R.G. (1998), "Capital-Market Imperfections and Investment", Journal of Economic Literature Vol. XXXVI (March 1998), pp. 193-225.
- Hubbard, R.G., Kashyap, A.K. and Whited, M. Toni, (1993), "Internal Finance and Firm Investment", NBER Working Paper No. 4392.
- Jaffee, D. and Russell, T., (1976), "Imperfect Information, Uncertainty, and Credit Rationing", The Quarterly Journal of Economics, Vol. 90, No.4. pp. 651-666.
- Jayaratne, J. and Wolken, J. (1999), "How important are small banks to small business lending?: New evidence from a survey of small firms", Journal of Banking & Finance, Volume 23, Issues 2-4, February 1999, Pages 427-458.
- Jovanovic, Bojan (1982) "Selection and the Evolution of Industry" Econometrica, Vol. 50, No. 3. (May, 1982), pp. 649-670.
- Kaplan, S.N., and L. Zingales (1997), "Do investment-cash flow sensitivities provide useful measures of financing constraints?", Quarterly Journal of Economics 112:159–216.
- Kaplan, S.N., and L. Zingales (2000), "Investment-cash flow sensitivities are not valid measures of financing constraints", Quarterly Journal of Economics 115:707–712.
- Kim, M. Kristiansen, E.G. and Vale, B. "Life-Cycle Patterns of Interest Rate Markups in Small Firm Finance", Working Paper August 11, 2008
- Klemperer P, (1987), "Markets with Consumer Switching Costs", The Quarterly Journal of Economics, Vol. 102, No. 2 (May, 1987), pp. 375-394.
- Laeven, L., (2003). Does financial liberalization reduce financing constraints? Financial Management 32 (1), 5-34.
- Levine, R. (2005) "Finance and growth: Theory and evidence", Handbook of economic growth, 2005.
- Li, W. and Weinberg, J., "Firm-specific Learning and the Investment Behavior of Large and Small Firms". International Economic Review, Vol. 44, No. 2, May 2003.
- Love, I., (2003). Financial development and financing constraints: international evidence from the structural investment model. Review of Financial Studies 16 (3), 765-791.
- Lyandres, E. (2007), "Costly external financing, investment timing, and investment–cash flow sensitivity". Journal of Corporate Finance 13 (2007) 959–980.
- Marquez, Robert (2002). "Competition, Adverse Selection, and Information Dispersion in the Banking Industry." The Review of Financial Studies 15, 901-926.
- Modigliani, F. and Merton, H.M., (1958), "The Cost of Capital, Corporation Finance and the Theory of Investment", The American Economic Review, Vol. 48, No. 3 (Jun., 1958), pp. 261-297.
- Moulton, B. (1990), "An ilustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units", The Review of Economics and Statistics, Vol. 72 No. 2 (May, 1990), 334-338.
- Moyen, N. (2004), "Investment-Cash Flow Sensitivities: Constrained versus Unconstrained Firms". The Journal of Finance, Vol LIX, No. 5, October 2004.
- Myers, S.C., and N.C. Majluf (1984), "Corporate financing and investment decisions when firms have information that investors do not have", Journal of Financial Economics 13:187–222.

- O’Hara, M. (1997), “Market Microstructure Theory”, Blackwell Publishing Ltd.
- Oliner, S.D., Rudebusch, G.D., 1992. Sources of the financing hierarchy for business investment. Review of Economics and Statistics 74 (2), 643-654.
- Ongena, S. and Smith, D. (2000), `What determines the number of bank relationships? Cross-country evidence', Journal of Financial Intermediation 9, 26-56.
- Ongena, S. and Smith, D. (2000), “Bank relationships: A survey”, In P. Harker and S.A. Zenios (eds), The Performance of Financial Institutions, Cambridge University Press .
- Pagano, M. and Jappelli, T. (1993). “Information Sharing in Credit Markets”. Journal of Finance 43:1693-1718.
- Pagano, Marco (1993). "Financial Markets and Growth. An Overview." European Economic Review 37, 613-622.
- Petersen M., G. Rajan R.G., (1995) “The Effect of Credit Market Competition on Lending Relationships” Quarterly Journal of Economics, 1995.
- Petersen M., G. Rajan R.G., (2002) “Does Distance Still Matter? The Information Revolution in Small Business Lending” The Journal of Finance, Vol. LVII, No. 6 • December 2002.
- Petersen, M.A, and Rajan, R.G. (1994). “The Benefits of Lending Relationships: Evidence from the Small Business Data”. Journal of Finance 49:3-37, 1994.
- Rajan, R. (1992): “Insiders and outsiders: The choice between relationship and arms length debt”, Journal of Finance 47, pp. 1367-1400.
- Rajan, R. and Zingales, L. (1998). “Financial dependence and Growth”, American Economic Review 88, 559-587.
- Schiantarelli, F. (1996). “Financial Constraints and Investment: Methodological Issues and International Evidence” Oxford Review of Economic Policy 12(2):70-89.
- Scott, JA and Dunkelberg WC (2005), “Competition for small firm banking business: Bank actions versus market structure”, Journal of Banking & Finance, Volume 34, Issue 11, November 2010.
- Shaikh, J M (2004) “Measuring and reporting of Intellectual Capital Performance Analysis” Journal of American Academy of Business 4(1/2), 439-448.
- Sharpe, S. A. (1990), “Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships”, Journal of Finance 45(4), 1069-87.
- Sharpe, SA. (1990), “Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships”, The Journal of Finance, Vol. 45, No. 4 (Sep., 1990), pp. 1069-1087.
- Stein, J.C. (2003), “Agency, Information and Corporate Investment” Handbook of the Economics and Finance, 2003, Elsevier Science B.V.
- Stiglitz, J., Weiss, A., (1981) “Credit Rationing in Markets with Imperfect Information”, The American Economic Review, Vol. 171, No. 3 (Jun 1981), 393-410.
- Stiglitz, J., Weiss, A., (1983), “Incentive Effects of Terminations: Applications to the Credit and Labor Markets”. The American Economic Review Vol. 73, No. 5 (Dec., 1983), pp. 912-927.
- Tobin, J. (1969), “A General Equilibrium Approach to Monetary Theory”, Journal of Money, Credit and Banking, Vol. 1, No. 1 (Feb., 1969), pp. 15-29.
- Townsend, R. (1979), “Optimal contracts and competitive markets with costly state verification” - Journal of Economic Theory, 1979

- Weinberg, J.A. (1994), "Firm Size, Finance and Investment" Federal Reserve Bank of Richmond Economic Quarterly Volume 80/1 Winter 1994.
- Weinstein, D. E. and Yafeh, Y. (1998), "On the costs of a bank-centered Financial system: Evidence from the changing main bank relations in Japan", Journal of Finance 53(2), 635-672.
- Whited, T., 1992. Debt, liquidity constraints, and corporate investment: evidence from panel data. Journal of Finance, 47 (4), 1425-1460.
- Whited, T.M. and Wu, G. (2006), "Financial Constraint Risk", Review of Financial Studies, Volume19, Issue 2, Pp. 531-559.
- Williamson, S.D. (1986). "Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing". Journal of Monetary Economics 18:159-179, 1986.
- Williamson, S.D. (1987). "Costly Monitoring, Loan Contracts, and Equilibrium Credit Rationing". Quarterly Journal of Economics 102:35-145, 1987.
- Wilson, P. F. (1993). The pricing of loans in a bank-borrower relationship. Working Paper, Indiana University.
- Winker, P. "Causes and Effects of Financing Constraints at the Firm Level", Small Business Economics 12, 169-181, 1999, Kluwer Academic Publishers.
- Yafeh, Y. and Yosha, O. (2001), "Industrial organization of Financial systems and strategic use of relationship banking", European Finance Review 5(1).
- Yan, Ying (1996) "Credit Rationing, Bankruptcy Cost, and Optimal Debt Contract for Small Business" Federal Reserve Bank of Cleveland, November 14, 1996.
- Yosha, O. (1995), "Information disclosure costs and the choice of Financing source", Journal of Financial Intermediation 4(1), 3-20.
- Zarutskie, R. (2004), "New Evidence on Bank Competition, Firm Borrowing and Firm Performance", Fuqua School of Business, Duke University, Mimeo, 2004.
- Zellner, A. (1971), "An introduction to Bayesian inference in econometrics", R.E. Krieger Pub. Co. (Malabar, Fla.) , ISBN 0898749956.