No Bank, One Bank, Several Banks: Does It Matter for Investment?[♣]

Alexander Karaivanov
Simon Fraser University

Sonia Ruano
Bank of Spain

Jesús Saurina
Bank of Spain

Robert Townsend
MIT

November, 2010

Abstract

This paper examines whether financial constraints affect investment decisions for older (larger) firms. We combine data from the Spanish Mercantile Registry and the Bank of Spain Credit Registry (CIR) to classify firms according to their number of banking relations: one, several, or none. Our empirical strategy combines two approaches based on a common dynamic model of firm finance and investment. First, using a standard Euler equation adjustment cost approach, we find that banked firms in our sample are likely to exhibit cash flow sensitivity while unbanked firms are not. Second, using structural maximum likelihood estimation, we find that unbanked firms' investment behaviour is close to that predicted by a model of (contingent) credit subject to moral hazard from unobserved effort, while single-banked and multiple-banked firms behave as if operating in a more limited financial environment, as in a traditional debt model. Firms in the unbanked category do not rely on bonds, equity or formal financial markets, but rather on other firms in a financial or family-tied group. To the best of our knowledge, we are among the first to document the importance of such groups in a European country. We control for reverse causality by treating bank relationships as endogenous and/or by appropriate stratifications of the relatively large sample.

JEL classification: C61, D82, D92, G21, G30

Keywords: financial constraints, bank lending, investment Euler equations, moral hazard, structural estimation and testing.

[♣] Research support from The US National Science Foundation, The National Institute of Health, The John Templeton Foundation, The Gates Foundation through the University of Chicago Consortium on Financial Systems and Poverty and The Social Sciences and Humanities Research Council of Canada is gratefully acknowledged. This paper is the sole responsibility of its authors and the views presented here do not necessarily reflect those of Banco de España or the Eurosystem. We thank Vicente Salas and Kenichi Ueda for comments and suggestions on earlier versions of this paper. We also appreciate the excellent comments from the Editor and an anonymous referee of the Banco de España Working Paper Series, and comments from the participants in the Far East Meeting of the Econometric Society, the Annual Congress of the European Economic Association, the Society of Economic Dynamics Annual Meeting and the Econometric Society World Congress. Any remaining errors are our own.
Address for correspondence: Sonia Ruano; C/ Alcalá 48, 28014 Madrid, Spain. Tlf: + 34 91 338 62 39; e-mail: sonia.ruano@bde.es.

1. Introduction

Over the past two decades the bulk of the literature on investment, at the micro level, has focused on the influence of financial constraints on firms' investment decisions.[1] Theorists have used models of information and incentive problems in capital markets to motivate the role and origin of financial constraints. Both adverse selection, due to asymmetric information between the borrower and potential lenders about the quality of the investment project, and moral hazard, due to the costly monitoring of managers' allocation of resources, can introduce a wedge between the cost of internal and external finance.[2] In both cases the ultimate result is limited access to external finance, so that firms have to restrict investment when internal cash is insufficient to invest at the first-best level.[3]

The empirical literature has documented that, in addition to the fundamentals determining investment dynamics (e.g., expected future profitability and the user cost of capital), firms' internal funds affect in a positive and significant manner their investment decisions, with more intensity for firms affected by capital market imperfections. Since the seminal paper of Fazzari, Hubbard and Petersen (1988), an extensively used empirical strategy consists of separating "constrained" and "unconstrained" firms according to *a priori* assumptions on the likelihood that firms are subject to information or incentive problems and then testing whether the neoclassical investment equation derived under the assumption of perfect capital markets with adjustment costs describes the behaviour of financially unconstrained firms but not that of financially constrained firms.[4] Typical firm characteristics used in the stratification approach as *a priori* proxies of the importance of capital market imperfections are firm's size, age, and dividend policy.[5] A few stratify by relationship with industrial or financial groups, and this turns out to be related to our findings.[6]

In parallel, a literature on relationship banking has developed based on the premise that information about borrowers is extremely important in the lending process, especially for small firms who are considered "informationally opaque".[7] Banks are thought to have a comparative advantage in

[1] Schiantarelly (1996) and Hubbard (1998) provide extensive reviews of the literature.
[2] On adverse selection see Jaffee and Russell (1976) or Stiglitz and Weis (1981). Myers and Majluf (1984) focus on information problems affecting equity financing. The effects of moral hazard are treated in Jensen and Meckling (1976) among others. Williamson (1987) derives the possibility of credit rationing in the context of optimally designged contracts under the assumption that profit outcomes can only be observed at a cost. On costly state verification see also Townsend (1979) and Diamond (1984).
[3] The gap between the cost of external debt and of internal funds can be explained by information costs but it could also be due to tax factors or other transaction costs.
[4] There are two main empirical approaches, pioneered by Fazzari, Hubbard and Petersen (1988) and Bond and Meghir (1994) respectively. The former approach is based on the estimation of the Q model of investment, extended by including a proxy for the availability of internal funds and testing for differences in the sensitivity of investment to cash flow between "constrained" and "unconstrained" firms. The latter approach, which we use here, estimates the investment Euler equation which should be valid for unconstrained firms but mis-specified for constrained firms. In both cases, the investment equations are derived from the neoclassical model under the assumptions of perfect competition, constant returns to scale, and quadratic adjustment costs.
[5] Schiantarelli (1996) provides an exhaustive list of the criteria used in the literature to partition firms according to their likelihood of being subject to financial constraints and discusses the empirical literature for every criterion.
[6] Although we do not stratify by relationship with industrial or financial groups, we find that a sizeable number of older, larger firms are unbanked. A major source of funds for these firms is debt with other enterprises in a financial group. Persistently unbanked firms tend to be family-tied (in either pyramidal or informal structures). More on this below.
[7] For comprehensive reviews of the literature on relationship banking see Boot (2000) and Elyasiani and Goldberg (2004).

producing information about borrowing firms (Diamond, 1984; Bond, 2004). To the extent that bank relationships mitigate asymmetric information problems between creditors and borrowers and facilitate monitoring, the incidence of financial constraints on firms' investment decisions and, consequently, the sensitivity of investment to cash flow, could differ across groups of firms stratified by the strength of their banking relationships.[8]

Although empirical research evaluating the costs and benefits of banking relationships has been extensive, papers relating differences in firms' investment behaviour with bank-firm relationships are relatively scarce.[9] In most cases, the results suggest that close bank relationships are associated with reduction in the sensitivity of investment to cash flow (Elston, 1996, for Germany; García-Marco and Ocaña, 1999, for Spain;[10] and Houston and James, 2001, for the USA). In one case the authors find no significant effect of bank relationships (Fuss and Vermeulen, 2006, for Belgium). In another case the result is a positive effect on cash flow sensitivity (Fohlin, 1998, for Germany during the formative years of universal banking).[11]

We use data from the Spanish Mercantile Registry and the Bank of Spain Credit Registry (CIR) to classify firms according to whether they have borrowed from a single Spanish credit institution, from several institutions, or from none. We restrict our attention to firms that are at least ten years old and not listed on the stock exchange. We find that a significant number of those firms (more than 3,500 firms every year, or 10% of our sample) are unbanked, that is, they do not borrow from any bank (domestic or foreign branch) in Spain. These firms' debt is non-trivial, but it comes in large part from trade credit and from other firms which are members of a group of family-related firms. Banked firms are in turn divided into single-banked and multiple-banked according to the number of banks from which they borrow. We examine whether there are financial constraints for each of these types of firms using two related but distinct empirical approaches based on a common dynamic model of firm finance and investment.

Our first approach is based on the Euler equation adjustment cost approach to investment (Bond and Meghir, 1994). We find that single- and multiple-banked firms are most likely subject to financial

[8] The strength of firms' banking relationship has been defined in the literature as: the relationship's *length* (Petersen and Rajan, 1994; Harhoff and Körting, 1998; Berger and Udell, 1995 and Degryse and Van Cayseele, 2000), the relationship's *scope*, i.e., type and number of financial services (Petersen and Rajan, 1994 and Degryse and van Cayseele, 2000), bank's ownership (Elston, 1996; García-Marco and Ocaña, 1999; Fohlin, 1998 and Chirinko and Elston, 2006;) and the *number of banks* (Petersen and Rajan, 1994; Harhoff and Körting, 1998; Houston and James, 2001 and Fuss and Vermeulen, 2006).

[9] With the exception of Houston and James (2001) and Fuss and Vermeulen (2006) who differentiate between single and multiple banked firms, the rest of the papers only consider the presence of banks as shareholders or on firms' boards of directors.

[10] Carbó, Rodríguez and Udell (2008) use a disequilibrium model to classify Spanish firms as 'constrained' and 'unconstrained' and find that investment is predicted by bank loans for unconstrained firms and by trade credit for constrained firms. Although not specifically focusing on the influence of banking relationships, the influence of financial constraints on investment decisions for Spanish firms is also analysed in Alonso-Borrego (1994), Estrada and Vallés (1998) and Hernando and Tiorno (2002).

[11] Other papers investigate differences in the role of bank relationships in alleviating financial constraints at different stages of the credit cycle. As an example, see Vickery (2005) who presents evidence that the benefits of bank relationships are higher when credit conditions are poor or deteriorating. Similarly, Gan (2007) investigate whether or not bank liquidity shocks reduce firm investment, paying special attention to the role of bank relationships in the propagation of downturns.

constraints. Unbanked firms are not. Our results are consistent with the literature which argues that bond markets are accessible only for larger firms. Yet in Spain higher levels of finance beyond bank borrowing comes not from bonds but from trade credit and the Spanish analogue to *chaebols* or *zaibatsu* business groups (Hoshi, Kashyap and Scharfstein, 1990). Though these relationships are not necessarily official or formal in many cases, we find evidence of the importance of family ties.[12] Specifically, we find that firms which persistently finance their investment projects without relying on banks, bonds, or other financial markets tend to be linked to other firms in formal or informal family groups.[13] We draw this conclusion from several case studies in which we carefully reconstruct the cross-firm relationships for several large persistently unbanked firms. Although the importance of family-related firms has been previously studied,[14] we believe we are among the first to document the importance of unbanked financial groups in a European country.[15]

Naturally, banking relationships, or the lack thereof, are likely to be endogenous and it is difficult to establish the direction of causality. Firms that are unconstrained may be happy within business groups as they do not need to worry about financing, for example. The group may not be playing an active role in alleviating constraints. Likewise, constrained firms may be precisely those that are unable to participate in a group and thus are likely to go to a bank. In that sense, limited observed finance by single banks *per se* may not be the cause of the constraint. The Bond and Meghir approach allows the econometrician to control for the endogeneity of bank relationships using appropriate lags of this variable as instruments in the GMM estimation.

In our second approach, we estimate our dynamic firm investment model under four financial regimes (autarky, non-contingent debt, moral hazard constrained credit and complete markets), using simulated maximum likelihood. The advantage of this approach is that we can get inside the "black box" of the credit relationship. Indeed, we first see how far we can get without imposing adjustment costs, so that all action is in the explicitly micro-modelled financial or information constraints. In effect, we combine theoretical dynamic models of credit constraints in investment with applied empirical work. We then allow for both adjustment costs and distinct financial regimes. Without adjustment costs we find that risk-neutral unbanked firms have a financial structure that is close to

[12] Other authors have also suggested that group membership relaxes financial constraints, see Schiantarelli (1996) for a review and Samphantharak (2003) for evidence from Thailand. Others (e.g., Bertrand, Mehta and Mullainathan, 2002) have recognised the existence of "tunnelling" (transfer of resources across firms in a pyramidal structure from a lower level firm to a higher level firm) and "propping" (transfer of resources in the opposite direction) and their impact on the relationship between controlling and minority shareholders .

[13] We label *formal* or *pyramidal* those structures in which firms are linked through shareholder/subsidiary relationships. We label *informal* relations those cases in which several firms, not linked through shareholder/subsidiary relationships, are related through family ties and/or other "coincidences" such as, common addresses, telephone numbers, etc.). In general, we consider two firms family-tied if they have one or more persons of the same family in their boards or lists of shareholders. On the nature of the agency problems within firm pyramids see Bertrand and Mullainathan (2003).

[14] Bertrand and Schoar (2006) provide a survey of world-wide evidence on family controlled businesses. They contrast theoretical arguments for the higher efficiency of family firms vs. 'cultural' theories claiming the contrary. More generally, Khanna an Jafeh (2007) review business groups with particular attention to family groups.

[15] The literature on family-tied firms in European countries includes Cronqvist and Nilsson (2003) for Sweden, Sraer and Thesmar (2004) for France and Maury (2006) for a sample of western European countries, among others. See also Morck (2005) for an exhaustive list of studies on the history of business groups in Germany, Italy, the Nethelands, Sweden, UK and other countries.

that predicted by a model of moral hazard with unobserved effort (complete markets / full information comes second), while risk-neutral single- and multiple-banked firms have financial structure which is more limited, essentially as in a traditional non-contingent debt model. The debt model also fits best the whole sample data. Allowing for quadratic capital adjustment costs in each of the four regimes weakly improves their fit with the data but, reassuringly, the best-fitting regime in each sub-sample defined by firms' number of banking relations remains the same as in the estimations without adjustment costs. Finally, these conclusions are robust to allowing for risk aversion.

To control for the possible endogeneity of bank relationships in this second approach we feature firms that have been continuously in only one category for the entire time period. For these continuing-status firms, the dynamic programming problem implies that they would behave approximately as if they were set to be in that category forever, especially if at the end of the sample they are still not near making a transition. The latter assumption cannot be tested, however. Also, firms in one category of bank relationships may be substantially different from those in another category for other reasons, and such heterogeneity is not yet incorporated into our model. But at least we follow literally the dictates of the theory that does allow for some (if not all) heterogeneity.

Though sensitivity of investment to cash flow has been widely interpreted as evidence of financial constraints affecting firms' investment decisions, the literature is not without controversy. As argued by Kaplan and Zingales (1997, 2000),[16] a firm is either constrained or unconstrained, a binary variable, and thus the degree of cash flow sensitivity does not necessarily indicate the severity of the financial constraint. Recent work by Gomes (2001), Cooper and Ejarque (2003) and others have challenged further the Fazzari et al. (1988) interpretation by demonstrating that financing constraints are neither necessary nor sufficient for cash flow effects.[17] On the other hand, Schmid (2009) shows that, in a model with incomplete markets due to limited commitment, the cash flow term is significant and therefore suitably modelled financial frictions can indeed rationalize the evidence on cash flow sensitivity.

Our results contribute to this debate given that we look at both investment sensitivity to cash flow and the estimation of the underlying financial regime. Our structural estimation results for groups of firms stratified by their number of banking relations reinforce the interpretation that the cash flow sensitivity of single-banked firms is associated with stricter financial constraints on investment decisions. To our knowledge, this combination of approaches is unique to our paper and provides consistent evidence of financial market imperfections for banked firms.

[16] For a response to this critique, see Fazzari, Hubbard and Petersen (2000).

[17] Papers that reject the monotonic relationship between cash flow sensitivity and financial constraints include Fohlin (1998), Cleary (1999), Houston and James (2001), and Fuss and Vermeulen (2006).

Our empirical findings are in line with those of Fohlin (1998) who documents that investment is less sensitive to internal resources in unbanked firms, contrary to what is reported in the rest of the papers for other countries, including Spain. One reason is that most authors compare banked and unbanked firms defined according to the presence of banks in the list of shareholders or on the board of directors. This is a valid but, in our opinion, restrictive way of distinguishing banked vs. unbanked firms. Instead, we define unbanked firms simply as those not borrowing from any bank. In contrast to the few papers which define bank relationships as we do, we also study unbanked firms in addition to comparing across single- and multiple-banked firms.[18]

With respect to previous studies on Spanish firms (García-Marco and Ocaña, 1999), any discrepancies in the results are likely due to differences in the definition of bank relationships, differences in the size and composition of their sample (limited to a few large and listed manufacturing firms), and/or differences in the sample period.

One substantial difference we have with the previous literature is that we restrict attention to firms that are at least ten years old. Our sample thus contains unbanked firms that are older and larger than small start-ups and young firms without credit history, possibly unbanked for that reason. The latter group is *a priori* likely to be more constrained than any other group, and we do not want to mix such firms in our sample of unbanked firms, which may be such because they have other sources of finance and choose not to borrow.

2. The model

The firm maximizes a 'utility' function, $u(c,z)$ where c is dividends to the owner, analogous to consumption, and z is labour effort in management and production, as if the firm were run by a single person or were a conglomerate with perfect markets within so that we have Gorman aggregation. The utility $u(c,z)$ is separable in c and z and we allow it to be linear in c to accommodate risk neutrality.

The firm's production function maps current effort, z and capital, k into output, y. The general notation is $y = F(k,z,\theta^q)$, where θ^q is a firm-specific productivity shock (superscripts are the way we distinguish this shock from others below). The firm's *net profits*, $q(k,z,\theta^q)$ equals revenues (output sales) less expenses for hired labour and other material costs, which we do not write explicitly. This general setup allows the shock θ^q to be auto-correlated, and this is important in the

[18] On the determinants of the number of bank relationships see Ongena and Smith (2000) and Guiso and Minetti (2007).

cash flow versus productivity debate, but we focus the exposition here on the case where shocks are i.i.d. over time, and over firms.

Capital depreciates at rate δ so we have the usual law of motion, $k_t = (1-\delta)k_{t-1} + I_t$, with I_t denoting time-t investment. As in Bond and Meghir (1994), investment I_t immediately adds to the capital k_t used in the current period.[19] There is a marginal cost, $g \geq 0$ of adjusting the current capital level. It is subtracted from firm's current net profits q_t. Using the standard quadratic form assumption, we write:

$$g(k_t, I_t; \theta_t^I) = \frac{1}{2} b k_t \left[\frac{I_t}{k_t} - \varsigma - \theta_t^I - b_j \right]^2$$

Here, θ_t^I is a shock to the investment adjustment cost, b_j is a firm-j specific term, and b and ς are parameters to be calibrated or estimated.

The firm discounts future profits with factor β. For the most part we take β as the inverse of the market rate for borrowing and lending, $1/R$, where $R = 1 + r$ is the gross rate of interest, but more generally, we allow β and $1/R$ to differ, for example, $R \geq 1/\beta$, so that the firm can achieve higher utility in the future by saving at the market rate. We can also easily introduce two rates, one (R^B) for borrowing and the other (R^S) for saving, with $R^B \geq R^S$. Finally, let the schedule of borrowing rates depend on the amount borrowed B through some function $\Psi(B)$.

To gain intuition into the mechanics of the model, consider first a two-period, t_0 and t_0+1, specialized version with risk neutrality and effort supplied inelastically at $z = \bar{z}$. Initial capital k_{t_0-1}, savings S_{t_0-1}, and debt B_{t_0-1} are given. The firm's problem is:

$$\max_{S_{t_0}, B_{t_0}, k_{t_0}, k_{t_0+1}, c_{t_0}, c_{t_0+1}} \quad c_{t_0} + \beta E_{t_0}(c_{t_0+1})$$

subject to:

$$\begin{aligned} c_{t_0} = \;& k_{t_0-1}(1-\delta) + q(k_{t_0}, z, \theta_{t_0}^q) - k_{t_0} - R^B B_{t_0-1} + B_{t_0} + \\ & + R^S S_{t_0-1} - S_{t_0} - \Psi(B_{t_0}) - g(k_{t_0}, I_{t_0}, \theta_{t_0}^I) \end{aligned}$$

$$c_{t_0+1} = k_{t_0}(1-\delta) + q(k_{t_0+1}, \bar{z}, \theta_{t_0+1}^q) - R^B B_{t_0} + R^S S_{t_0} - g(k_{t_0+1}, 0, \theta_{t_0+1}^I)$$

where S_{t_0+1}, B_{t_0+1} are zero and $c_{t_0} \geq 0$, $c_{t_0+1} \geq 0$, $S_{t_0} \geq 0$, $B_{t_0} \geq 0$.

---

19 In Bond and Meghir (1994) investment is subtracted as a cost in the net profit function. Note also that our q, as net profits, is already net of wages and other costs. In Bond and Meghir (1994) these non-capital input decisions are explicit.

Suppose a firm starts with a sufficiently low level of capital k_{t_0-1}, so that the marginal product of capital net of the adjustment costs at t_0+1 is sufficiently high that,

$$E\frac{\partial q\left(k_{t_0+1},\bar{z},\theta^q_{t_0+1}\right)}{\partial k_{t_0+1}}-E\frac{\partial g\left(k_{t_0+1},0,\theta^I_{t_0+1}\right)}{\partial k_{t_0+1}}>R^B\geq R^S\geq\frac{1}{\beta}.$$

Such a firm would like to invest at t_0, but is constrained by current-period resources. Thus, $c_{t_0}=0$ and $c_{t_0+1}\geq 0$. The constraint that dividends at t_0 be non-negative is binding, with a positive Lagrange multiplier $\lambda_{t_0}>0$. To augment resources for investing further, the firm will borrow up to the point that the expected net marginal product of capital is equated to the marginal cost of borrowing,

$$E\frac{\partial q\left(k_{t_0+1},\bar{z},\theta^q_{t_0+1}\right)}{\partial k_{t_0+1}}-E\frac{\partial g\left(k_{t_0+1},0,\theta^I_{t_0+1}\right)}{\partial k_{t_0+1}}=\frac{R^B}{1-\Psi'\left(B_{t_0}\right)}.$$

With increasing marginal borrowing costs, the tendency is for the firm to remain small and grow slowly.

On the other hand, suppose a firm starts with relatively high level of capital,[20] e.g., at the steady state value,

$$E\frac{\partial q\left(k_{t_0+1},\bar{z},\theta^q_{t_0+1}\right)}{\partial k_{t_0+1}}-E\frac{\partial g\left(k_{t_0+1},0,\theta^I_{t_0+1}\right)}{\partial k_{t_0+1}}=R^B\geq R^S.$$

Such a firm only invests to replace depreciated capital and does not borrow. If $R^B=R^S$, this firm is also indifferent about the path of dividends, thus $\lambda_{t_0}=0$. These large and older firms would show up as unconstrained.

Bond and Meghir (1994) establish more generally that firm's investment is sensitive to cash flow, q precisely when $\lambda_{t_0}>0$. They also allow for default on debt with a bankruptcy cost, a variety of taxes on dividends received by owners, and advantages for interest paid to lenders and for capital gains. Here, we follow Bond et al. (2003) assuming that there are no cash flow constraints or tax considerations and test empirically whether the unconstrained benchmark with zero Lagrange multiplier on dividends is rejected against the alternative.

Formally, let $\Pi(k_t,I_t,\bar{z},\theta^q_t,\theta^I_t)\equiv q(k_t,\bar{z},\theta^q_t)-g(k_t,I_t,\theta^I_t)$. The firm maximizes expected present discounted value at time t of current and future cash flows,

[20] The adjustment cost function g is increasing in I_{t0} which implies that the left hand side is increasing in k_{t0}.

$$\max_{\{k_{t+j}\}} E_t\left(\sum_{j=0}^{\infty} \beta^j \Pi(k_{t+j}, I_{t+j}, \bar{z}, \theta^q_{t+j}, \theta^I_{t+j}) \right)$$

$$\text{subject to: } k_{t+j} = (1-\delta)k_{t+j-1} + I_{t+j} \quad \text{for all } j.$$

This implies the following Euler equation characterizing the optimal investment path for any $t_0 \geq t$,

$$-\left(\frac{\partial \Pi}{\partial I}\right)_{t_0} = -(1-\delta)\beta E_{t_0}\left(\frac{\partial \Pi}{\partial I}\right)_{t_0+1} + \left(\frac{\partial \Pi}{\partial k}\right)_{t_0}.$$

Assuming constant returns to scale (inclusive of labour and other inputs), competitive markets, and exploiting the quadratic form of the adjustment costs function *g*, we obtain,

$$\left(\frac{I}{k}\right)_{t_0} - \alpha_1\left(\frac{I}{k}\right)^2_{t_0} = \alpha_2 E_t\left(\frac{I}{k}\right)_{t_0+1} + \alpha_3\left(\frac{q}{k}\right)_{t_0} + \alpha_0$$

Current investment is positively related to expected future investment and to current average profits reflecting the marginal productivity of capital. Thus, all expectation terms are captured by a one-step-ahead forecast.

To implement this equation empirically, replace the expectation by its realised value $\left(\frac{I}{k}\right)_{t_0+1}$ plus a forecast error, move this term to the left-hand side and backdate. Adding the subscript *j* for firm-specific variables and generalising the time index to *t*, we obtain the following empirical specification,

$$\left(\frac{I}{k}\right)_{jt} = \beta_1\left(\frac{I}{k}\right)_{jt-1} + \beta_2\left(\frac{I}{k}\right)^2_{jt-1} + \beta_3\left(\frac{q}{k}\right)_{jt-1} + \beta_4\left(\frac{y}{k}\right)_{jt-1} + d_t + \eta_j + \varepsilon_{jt} \qquad [1]$$

The output-to-capital ratio term, $\left(\frac{y}{k}\right)_{jt-1}$ is introduced to allow for either non-constant returns to scale or monopolistic competition in the product market. The term d_t captures common components in expectations, e.g., commonly observed aggregate shocks and η_j is a firm-specific effect. Equation [1] is the basic specification that we use in this paper. Under the null of no financial constraints, Bond et al. (2003) show that $\beta_1 \geq 1$, $\beta_2 \leq -1$, $\beta_3 < 0$ and (under constant returns to scale) $\beta_4 \geq 0$.

On the other hand, firms could be risk averse or even under risk neutrality there could be alternative credit market imperfections not captured by the model of limited debt with borrowing costs assumed above. As in Karaivanov and Townsend (2010), we allow for these possibilities by considering a range of financial 'regimes'. One extreme is *autarky* in which there is no borrowing and saving. A second financial regime we consider is *borrowing* without default (i.e., no risk contingencies of any kind).

Further, to study more complex credit regimes which allow for risk-contingent premia, transfers, and debt, we write the firm's objective function as

$$u(c,z)+\beta w',$$

where $c=\tau+(1-\delta)k-k'$ is current dividend and where, conforming with Bond and Meghir's (1994) timing, k denotes the beginning-of-period capital stock while k' denotes the capital stock actually used in production. In this formulation, it is as if all profits go to the lender but some part is then returned to the firm via the (output-contingent) transfer τ. The firm bears the cost for changes in the capital stock, i.e., it pays for investment but, assuming full observability, capital k and investment I are effectively under the control of the lender. The variable w' reflects next period promised utility, that is, discounted expected future utility. In the risk-neutral case it equals expected dividends from next period onward. Importantly, effort z may not be observed by the lender and so the firm must be given incentives to perform – a moral hazard problem.

Writing c as function of observed net profits, $c(q)$ as in a profit sharing arrangement, we have, for all k and all $\hat{z}\neq z$ the following incentive compatibility constraints,

$$\sum_{q,k'}\text{Prob}(q\mid z,k')[u(c(q),z)+\beta w'(q)]\geq\sum_{q,k'}\text{Prob}(q\mid \hat{z},k')[u(c(q),\hat{z})+\beta w'(q)] \qquad \text{(ICC)}$$

To compute solutions to this *moral hazard* credit regime we use the joint probability distribution over all possible allocations $\pi(\tau,q,z,k',w'\mid w,k)$ and re-write the problem as a dynamic linear program in the probabilities π (see Prescott and Townsend, 1984 and Karaivanov and Townsend, 2010 for more details). This very general way of writing the problem takes care of any potential problems with non-convexities and the first-order approach since it allows for ex-ante lotteries over allocations. Assume that all variables (τ, q, z, k', w') belong to discrete finite grids. With a large number (continuum) of firms, π will also be the frequency distribution we see in the data, if all variables were observed. In practice, we see only k, i and q so their joint distribution is obtained from the model by integrating out the unobservables. In terms of these joint probabilities, the moral hazard constraints (ICC) can be written $\forall\bar{z},\hat{z}$ as

$$\sum_{\tau,q,k',w'}\pi(\tau,q,\bar{z},k',w'\mid w,k)[u(\tau+(1-\delta)k-k',\bar{z})+\beta w']\geq\sum_{\tau,q,k',w'}\pi(\tau,q,\bar{z},k',w'\mid w,k)\frac{\text{Prob}(q\mid\hat{z},k')}{\text{Prob}(q\mid\bar{z},k')}[u(\tau+(1-\delta)k-k',\hat{z})+\beta w']$$

Here w denotes the promised utility in the entering period (the state variable capturing past history). This promise must be met, so for $\forall k$ we must have ("promise keeping"):

$$\sum_{\tau,q,z,k',w'}\pi(\tau,q,z,k',w'\mid w,k)[u(\tau+(1-\delta)k-k',z)+\beta w']=w$$

The lender (or collection of lenders) is modelled as a risk neutral principal maximizing discounted expected returns when facing a long-term contract with a firm with current promise *w* and initial capital *k*

$$V(w,k)=\sum_{\tau,q,z,k',w'}\pi(\tau,q,z,k',w'|\,w,k)\left[q-\tau+\frac{1}{R}V(w',k')\right]$$

This objective function is maximised subject to the promise keeping constraint, the incentive constraint (ICC) and, in addition, to the Bayes-rule compatibility constraints,

$$\sum_{\tau,\,w'}\pi(\tau,\,\bar{q},\,\bar{z},\,\bar{k}',\,w'|w,k)=\text{Prob}(\bar{q}\,|\,\bar{z},\,\bar{k}')\sum_{\tau,\,q,\,w'}\pi(\tau,\,q,\,\bar{z},\,\bar{k}',w'|w,\,k)\quad\text{for all }\ \bar{q},\bar{z},\bar{k}'$$

and an "adding-up" constraint for the joint probabilities:

$$\sum_{\tau,q,z,k',w'}\pi(\tau,\,q,\,z,\,k',w'|w,\,k\,)=1\,.$$

The linear programming approach has two important advantages. First, since the constraint set is convex by construction we avoid potential issues with non-convexities due to the incentive constraint. Second, the linear programs' solutions enter directly into the likelihood that we use to estimate the model. For consistency, we also compute the solutions to the financial autarky (no borrowing/lending/insurance) and more standard borrowing/saving regimes using the linear programming method. In our fourth and final regime we drop the incentive constraint on effort (ICC) from the problem above, which we call the *full information* (full insurance, complete markets) setting.

In sum, we set up four financial regimes ranging from most constrained (autarky) to least constrained (full information) in terms of availability of credit and, if firms are risk-averse, insurance. We estimate and test these four regimes and their implications for firms' investment behaviour and sensitivity to cash flow using maximum likelihood in Section 4.2 below. The results are then compared with the investment Euler equation approach.

Both our Euler equation and structural MLE approaches can be generalized further by making the inter-bank market less than perfect, to distinguish the situation of the lender. The more difficult it is for the lender to acquire funds and pass along risk, the less can be done for the borrower. Our methods also allow for modelling transitions across banking status. See Appendix 1 on how our model can be extended to capture this.

3. Data and descriptive analysis

3.1. The data on bank-firm relationships

Our overall data sample contains information drawn from two main micro data sets. The first source is SABI-INFORMA which provides economic and financial information reported annually by Spanish firms in their public financial statements deposited at the Spanish Mercantile Registry. In addition, SABI-INFORMA provides static[21] information on some firm characteristics such as location (province), age, and business activity (4-digit CNAE-93 code, 488 industries in total) of the firm. Additionally, firms are categorised into private/public and listed/unlisted.

The second data source we use is the Credit Register (CIR) of *Banco de España* (the Spanish Central Bank). This database is a census of loans granted by Spanish credit institutions (mainly, commercial banks, savings banks and credit cooperatives) or subsidiaries and branches of foreign banks operating in Spain to Spanish firms.[22]

The combination of these two data sets allows a categorization of firms according to their number of credit relationships with banks (those firms included in the SABI-INFORMA sample that have at least one loan registered in CIR) vs. firms that are unbanked (otherwise). In the case of banked firms, we further differentiate between single-banked and multiple-banked firms. The resulting database contains annual information on economic characteristics and banking relationships for a large sample of Spanish non-financial firms during the period 1992-2004. In the rest of the paper, we refer to this database as SABI-CIR.[23] It is an unbalanced panel, as a consequence of entry and exit of firms. These are due, first, to the natural processes of firm "births" and "deaths"; second, to changes in the composition of the firms' portfolio for which INFORMA collects information from the Spanish Mercantile Registry and, finally, to the filtering[24] we did on the raw data that may lead to dropping some, but not necessarily all, observations for a given firm. The complete database contains approximately 3,621,000 observations, corresponding to around 773,000 firms which are observed in different time intervals over the period 1992-2004.

Since government-owned firms' investment decisions could be taken according to different criteria than those used by private firms, we do not include firms owned by central, regional or local governments in our analysis. This means dropping around 1,200 firms representing 0.33% of the non-

---

21 Information is static in the sense that it reflects the status of the firm at the moment the information was drawn which does not necessarily coincide with the status of the firm over the years—though we do construct a panel and use it, below.

22 The Bank of Spain Credit Registry (CIR) contains monthly information on *all* credits over a minimum threshold granted by credit institutions operating in Spain to Spanish borrowers. Given that the minimum threshold has been very low over the sample period, especially after 1995 (6.000 Euros), this database can be considered as a census of the banking loans granted to non-financial firms. For detailed analysis of the database see the *Memoria de la Central de Información de Riesgos*, published by Banco de España since year 2005 as well as Jiménez et al (2006) and Jiménez et al (2009).

23 An exhaustive list of all variables contained in our dataset and their definitions is provided in Appendix 2.

24 We dropped, for any given year, observations corresponding to firms declaring interest payments equal or higher than total debt or negative equity. Moreover, we dropped the value of a variable if the variable is, by definition, non-negative, but the firm declares a negative value for it. To be precise, we apply this filter to: sales, total assets, tangible assets, financial income, financial expenses, short term debt, long term debt, commercial debt, and cash.

financial firms in the SABI-CIR sample. Additionally, we exclude publicly listed firms (around 165 firms) since their access to funds from capital equity markets as an alternative to bank loans would show as a different financing mechanism and could blur our overall message.

A priori, the properties of the unbalanced panel, characterised by a wide coverage of age, size and industry categories,[25] make it attractive for studying the existence of financing restrictions influencing investment patterns. Of particular interest is the role of banking relationships in alleviating, or not, financial constraints on Spanish non-financial firms which, as non-listed companies, tend to be informationally opaque. Nevertheless, in practice, initial results based on the entire panel, including young and small firms, were rather unsatisfactory. We were mixing two very distinct categories of unbanked firms, young start-ups with older, more established entities. Thus, we decided to concentrate on the latter, older group.

We selected those firms for which we have information on the dependent and the explanatory variables in equation [1] for all years between 1997 and 2004.[26] Consequently, the sample excludes start-ups entering after 1997. This selection procedure generates a balanced panel containing information on 44,644 firms observed each year between 1996 and 2004.[27] We further filtered the sample by eliminating firms younger than ten years old in 1997. [28]

In sum, the data set we use in this paper contains 410,882 observations corresponding to all firms in the SABI-CIR database with ten or more observations available for period 1993-2004 that are at least ten years old in 1997. From now on we refer to this sample as the "selective sample". Table 1 provides the distribution of firms in this sample according to the number of years in which the firm is present in the dataset. Firms in the sample represent, on average, 10.6% of all firms in the SABI-CIR database, covering around 25% of aggregate total assets, aggregate debt and aggregate banking debt. With respect to the whole population of Spanish non-financial firms, the coverage of our selective sample is, on average, around 16% of the aggregate total assets and around 17% of the total amount of loans given by Spanish credit institutions to Spanish non-financial firms.[29]

---

[25] In terms of age, 63% of total firms in the SABI-CIR sample are less than 10 years old. In terms of size, there is a high proportion of small firms (93% of the firms in the sample), where small firms are defined as those with total assets not higher than €5 million or total sales not exceeding €7 million.

[26] 1997 is the earliest year in which data is available (from the period 1993-1995) to estimate the investment equation [1] using 3- and/or 4-year lags of the variables on the right-hand side as instruments for the predetermined variables in the first-difference equation.

[27] We are aware that picking firms with complete data might introduce some selection bias because the probability of having complete data is likely higher for larger firms which possibly assign more resources to satisfy the informational requirements of the Mercantile Registry. We discuss this issue in detail in Appendix 3.

[28] We filtered by age all firms reporting this information (around 40% of the firms in the balanced panel) but also kept in our sample all firms that do not report their founding date. Subject to data availability, this way of filtering jointly with using a balanced panel, helps us exclude as many as possible young start-up firms.

[29] Commercial banks, savings banks, credit cooperatives and credit financial establishments. Credits to Spanish firms granted by branches and subsidiaries of foreign banks in Spain are also taken into account.

We investigate the association of banking relations with firm's investment decisions. For this purpose, we define and distinguish among three groups of firms depending on the number of banking relations maintained by the firm in a given year. The number of banking relations is defined as the number of banks reporting at least one loan to the Bank of Spain Credit Registry (CIR) for this firm and time period. According to this definition, in each year the firms in our sample can be classified into two categories: "unbanked firms", those with no banking loans registered in CIR, and "banked firms", those with at least one loan registered at CIR. Additionally, within the category of banked firms, we distinguish between firms with loans from a single bank only ("single-banked firms") and firms funded by two or more banks ("multiple-banked firms").

In addition, we distinguish firms according to their number of banking relations over the whole studied period (1997-2004). Specifically, we classify firms in the categories of "continuing unbanked", "continuing banked" and "switching-status". Within the group of continuing banked firms we also differentiate among "continuing single-banked" firms, "continuing multiple-banked" firms and "other continuing banked" firms (those that switch in terms of number of lenders).

3.2. Descriptive analysis

We first present descriptive statistics on the composition of our sample by firm type defined according to the number of banking relations in a given year and according to changes in this number over the studied period. We then examine the differences and similarities across unbanked, single-banked and multiple-banked firms in (i) firms' characteristics (age, size and industry); (ii) investment patterns; and, (iii) economic and financial ratios. At the end of this section we explore in more detail how unbanked firms finance their activities.

The year-by-year composition of the sample according to the firms' number of bank relations is shown in Table 2. Most firms (approximately 72%) maintain simultaneous relations with two or more different banks,. Within this category, more than 60% of firms have bank loans from no more than four banks.[30] Around 18% of all firms have a single bank relationship while around 10% of firms are unbanked.

This distribution of firms across these categories is stable over time and is very similar for firms similar in age. Table 3 provides the distribution of the number of banking relations for firms in different age groups, [10, 20), [20, 30), older than 30 years, and firms that do not report age. Only the

[30]The importance of the main bank in multiple banked firms is decreasing with the number of banking relations. In particular, the proportion of total loans from the main bank for the median multiple-banked firm goes from around 80%, for firms with two bank relations, to around 30%, for firms with ten or more relations.

category of medium-aged (20 to 30 years of age) firms differs slightly in composition from the rest with a somewhat higher fraction of multiple-banked firms (75.2%).

Firms' banking status is highly persistent over time, as shown in the transition matrices presented in Table 4. Persistence is remarkably high within the category of multiple-banked firms where 93.6% on average remain in the same category a year later. This percentage decreases only slightly to 86.4% when looking five years ahead. Firms that leave the category of multiple-banked firms are more likely to transition into the single-banked category than into the unbanked category.

Persistence is somewhat lower but still high in the category of unbanked firms for which 72.5% of the firms remain in this category one year later, around 63% two years later and almost one-half five years later. Approximately three-quarters of the firms exiting this category in a given year go to the single-banked category. The fraction of unbanked firms moving to the category of multiple-banked firms increases gradually with the time horizon from 8% to 21%.

Having a single bank relation, on average, is somewhat more transitory. Table 4 shows that the percentage of single-banked firms that remain in the same category in subsequent years varies from 68%, after one year to around 43% after five years. Approximately two-thirds of single-banked firms that switch to other categories move to multiple-banked status.

Table 5 reports the distribution of the number of years in which a firm is classified in the unbanked category over the period 1997-2004. The vast majority, 76.7% of all firms are "continuing banked", i.e. banked in all eight years. At the other extreme, only 2.4% of the firms in the sample are "continuously unbanked". Still, this number is statistically meaningful as these are almost 900 firms in absolute numbers.

Within the group of continuously banked firms, a large fraction remains multiple-banked over the entire period. Firms in this sub-category, labelled "continuing multiple-banked", represent 68% of all continuously banked firms. In contrast, firms remaining single-banked firms in all years (that is, "continuing single-banked" firms) represent only a small fraction (3.3%, 914 firms) of continuously banked firms. The residual sub-category, named "other continuing banked" firms, includes all continuing banked firms that change between single- to the multiple-banked status at least once in the sample period and represents around 29% of continuing banked firms.

Approximately 21% of firms switch their banked/unbanked status at least once between 1997 and 2004. As shown further in Table 6, within this category, most firms (more than 80%) change their banked/unbanked status once (43%) or twice (37%) in the sample period. Switching three times adds another 13%, four times - 6% and the rest of the numbers are negligible.

Firm age, size, industry

As anticipated from the earlier discussion, there are not many differences across bank status and age. In the more detailed breakdown given in Table 7 multiple-banked firms tend to be slightly older than unbanked and single-banked firms, as revealed by their median age of 19.5 (1.5 years higher than in the other two groups[31]) and by the higher dispersion in the upper tail of the age distribution. This suggests that the oldest firms in our sample tend to be concentrated in the category of multiple-banked firms. Overall, the firm age histograms are similar across the firm groups.

In contrast, in terms of size, measured by total assets, there are important differences among the three compared groups (Table 8). Multiple-banked firms are the largest, in terms of mean and median total assets followed by single-banked firms. Unbanked firms are the smallest. As shown in Table 8, the entire firm size distribution for multiple-banked firms is shifted to the right relative to that for single-banked firms, which in turn is shifted to the right relative to unbanked firms.

The distribution of firms across industries is largely similar for the three categories as shown in Table 9. Perhaps most remarkable is the presence of unbanked firms in all industries. Multiple-banked firms have higher relative weight in manufacturing and quarrying.

Investment patterns

Table 10 reports averages, over the period 1997-2004, of the firm-level distributions of relative investment, defined as the ratio of absolute investment to capital (for definitions of all variables see Appendix 2) for the whole sample and for the sub-samples defined by number of bank relations. Most remarkable is the wide heterogeneity in investment across firms within each category. In the whole sample, the median firm invests, on average, 18.5% of the value of its fixed assets. Firm investment ratios range from a minimum of -0.58,[32] to a maximum of 1.4. Based on banking status, the 25th, 50th, and 75th percentiles of the investment ratio are higher for multiple-banked firms than for the rest. Compared to single-banked firms, the investment ratio percentiles for unbanked firms are lower. These differences in investment could suggest that there may be distinctive technologies across firm categories and/or the influence of financial constraints affecting firms from different groups in a different manner. For example, unbanked firms may operate technologies that require less investment and hence demand less outside finance. This is, of course, an empirical question. When we stratify by

[31] Note that the age distribution in our selective sample is rightwards-biased due to the exclusion of new firms entering after 1996 and the exclusion of firms younger than ten in 1997. Still, the conclusions about the differences in age composition across unbanked, single-banked and multiple-banked firms that can be drawn considering the unbalanced panel are analogous to those for the selective balanced panel. In particular, the median firm is 8.3, 6.1 and 5.8 years old, in the categories of unbanked, single-banked and multiple-banked firms, respectively. There are also significant differences in the same direction in the bottom end of the distribution, e.g., the first age quartile equals 2 for unbanked, 2.9 for single-banked and 4.4 for multiple-banked firms.

[32] Investment is defined as the sum of the absolute change in physical assets and depreciation. Negative values of investment correspond to net sales of physical assets.

bank status in the maximum likelihood estimation in Section 4, we implicitly allow for different technologies across firms by allowing the technology parameter estimates to fit the data.

Regarding firms' gross investment (Table 11), the most remarkable feature is the difference in the fraction of firms with gross investment of zero across the three groups. Among unbanked firms this fraction reaches 11.5%, five percentage points higher than the fraction for single-banked firms and more than 9 percentage points higher than that for multiple-banked firms. Our model allows zero investment but does not distinguish investment events on the extensive margin (i.e., whether or not to invest at all) from the size of investment on the intensive margin, conditional on investing. A fixed cost rather than quadratic adjustment costs approach would distinguish between those cases but makes the Euler equations less tractable.

Additional autocorrelation analysis of relative investment (not reported in the tables) shows that investment decisions are more persistent for multiple-banked firms than for unbanked and single-banked firms. Moreover, there are small differences between unbanked and single-banked firms with changing sign depending on the time horizon (details are available upon request).

Economic and financial ratios

As shown in Table 12 (first row), the more banking relations a firm has, the higher its debt-to-assets (leverage) ratio. All percentiles of the distribution of the *leverage ratio* for multiple-banked firms are higher than those for single-banked firms, which in turn are higher than those for unbanked firms. The leverage ratio of the median multiple-banked firm is 66.4%, which is 15 and 29 percentage points higher than the median leverage ratio for single-banked and unbanked firms, respectively. Still, it is clear from the table that unbanked firms do obtain external funds (from non-bank creditors such as associated or affiliated companies or trade credit)[33] and in a significant amount. For half of these firms, debt varies between 20% (25th percentile) and 58% (75th percentile) of their assets. Thus, one should not think of unbanked firms as necessarily needing or receiving less credit, but rather as obtaining credit from alternative, non-bank sources. [34]

Single-banked and multiple-banked firms also exhibit differences in their *degree of bank dependence* defined as the ratio of banking debt to total debt (the second row in Table 12). In particular, multiple-banked firms tend to rely more heavily on banking debt than single-banked firms. The median value is 44.2% in the first group, more than twice that for the second group. However, the wide dispersion of

[33] *Debt with associated and affiliated companies* and, to a lower extent, *trade credit*, are the two main sources of funds for unbanked firms. For instance, in 1997 these two types of debt represent around 72% and 10% of unbanked firms' debt, respectively. This conclusion is drawn from the aggregated balance sheets built with data for the subsample of firms that report to the Spanish Commercial Register detailed information on the itemization of debt, which represent around 1% of the firms in our SABI-INFORMA database for 1997. The numbers are similar for other years.

[34] Commercial bonds are a marginal source of finance in Spain. For a detailed discussion on the debt composition of unbanked firms see Section 5.

this ratio in both groups indicates that firms within each group are highly heterogeneous in their dependence on banking debt as source of external funds.

The *liquidity ratio* (short term assets to short term liabilities, third row of Table 12) moves inversely with the intensity of banking relations. The Table suggests that unbanked firms tend to hold higher liquidity levels than firms in the other two categories, with a ratio of 43.8% for the median unbanked firm. Additionally, there are wide differences within banked firms. Liquidity is much lower for multiple-banked firms (median liquidity ratio around 8%) than for single-banked firms (median ratio above 23%).

There are no significant differences among the three groups in terms of the *cash flow to capital ratio*, a proxy for the availability of internal funds; the *sales to capital ratio*, which proxies for the productivity of firms' assets; and the *return-on-assets ratio (ROA)*, a standard measure of the profitability of the firm.

The last row of Table 12, presents the *default ratio* (on banking debt).[35] The default ratio is computed conditioning on the sub-sample of firms that default on their bank loans. The fractions of such firms are very small (below 1%) for both single- and multiple-banked firms. For the median firm, the default ratio is much higher for single-banked firms (100%) than for multiple-banked firms (35.3%).

4. Results

4.1. Estimation results based on the investment Euler equations

4.1.1. Empirical strategy and econometric issues

We first use the standard Euler equation approach to examine whether firms' investment behaviour and sensitivity to cash flow fluctuations vary with the firms' number of banking relations. For this purpose we estimate a version of the Euler equation [1] derived from the model of Bond and Meghir (1994) and also used in Bond et al (2003).

Bond and Meghir (1994) show that, under the presence of financial constraints, the basic Euler equation [1] is mis-specified. In particular, the predicted negative sign on the cash flow term is expected to fail. The explanation is that, in the presence of financial constraints, the cash flow term reflects, in addition to marginal productivity of capital, the influence of financing restrictions, manifested in a positive correlation between investment and cash flow.

35 The default ratio is the amount of non-performing bank loans expressed in terms of the total amount of bank loans.

In line with the literature, we adopt a sample splitting approach that uses the number of banking relations as a criterion to separate firms across categories that are presumably affected by financial constraints in a different manner. The parameters of the investment Euler equation are estimated for each group and then differences (if those exist) are tested.

The investment Euler Equation [1] is an autoregressive panel data model the consistent estimation of which requires employing appropriate panel data techniques.[36] Specifically, we use the first-differenced GMM estimator for autoregressive panel data models developed by Arellano and Bond (1991).[37] The Arellano-Bond method requires estimating a first-differenced transformation of Equation [1] to remove firm-specific effects. Because of the endogeneity of first-differenced $\left(\frac{I}{k}\right)_{jt-1}$ and $\left(\frac{I}{k}\right)^2_{jt-1}$ it is necessary to instrument these variables using appropriate lags of the dependent variable under the identifying assumption that the serial correlation in the error term ε_{jt-1} is limited.[38] We test the validity of our instruments by performing direct tests for serial correlation in the residuals.[39] As explained in more detail below, we find no evidence of serial correlation in the error term of order higher than two in the first-differenced equation. This suggests use of instruments dated *t-3* and earlier.

First-differencing Equation [1] implies a second source of endogeneity which also needs to be addressed. Namely, the other variables on the right hand side, $\left(\frac{q}{k}\right)_{jt-1}$ and $\left(\frac{y}{k}\right)_{jt-1}$ become endogenous even if they are assumed pre-determined. To deal with this problem, Arellano and Bond (1991) propose using lags of these additional explanatory variables to define additional moment conditions.

Finally, a firm's decision on its number of banking relationships can also be endogenous and co-determined with the decision on investment. One possibility is that the firm can adjust the number of banking relations according to the necessity of external funds to implement the optimal investment policy. Alternatively, the firm's investment decision may be limited by the number of banks which are

36 For a review of econometric methods for dynamic panel data models see Bond (2002).

37 As a robustness check we also re-estimated all specifications using the extended first-differenced GMM estimator proposed by Blundell and Bond (1998). See Appendix 4. In situations where first-differences are close to being innovations, the Arellano and Bond GMM estimator can be expected to perform poorly (bias and imprecision). Under further restrictions on the initial conditions process the Blundell-Bond estimator corrects this issue. However, as we will justify below, these stronger assumptions on the initial conditions appear not to be satisfied in our case and consequently the standard Arellano and Bond GMM estimator is our preferred option.

38 The order of lags valid as instruments depends on the autocorrelation structure of the error term. In general, if the error term has an MA(r) structure, then lags of order higher than r+1 for the dependent variable are valid instruments.

39 The null hypothesis of the autocorrelation test is the absence of autocorrelation of a given order.

willing to lend money to the firm. Whatever the direction, it is important to consider the potential endogeneity of the number of bank relations.

We treat this endogeneity in the following manner. First, we generalize the basic Euler Equation in [1] to allow for differences in the model parameters for the categories of unbanked and banked firms as in Equation [2],[40]

$$\left(\frac{I}{k}\right)_{jt} = \beta_1\left(\frac{I}{k}\right)_{jt-1} + \beta_2\left(\frac{I}{k}\right)^2_{jt-1} + \beta_3\left(\frac{q}{k}\right)_{jt-1} + \beta_4\left(\frac{y}{k}\right)_{jt-1} + d_t + \qquad [2]$$
$$+ \delta_1\left(\frac{I}{k}\right)_{jt-1} \times D0_{jt} + \delta_2\left(\frac{I}{k}\right)^2_{jt-1} \times D0_{jt} + \delta_3\left(\frac{q}{k}\right)_{jt-1} \times D0_{jt} + \delta_4\left(\frac{y}{k}\right)_{jt-1} \times D0_{jt} + d_t{\times}D0_{jt} + \eta_j + \varepsilon_{jt}.$$

where $D0_{jt}$ denotes a dummy variable which equals one if firm j is banked in year t, and zero otherwise. The parameters β_l, $l=1,...,4$, characterise the dynamics of investment in the reference group of unbanked firms, while the parameters δ_l, $l=1,..., 4$, quantify differences in β_l between the group of banked and unbanked firms. The year fixed effect term d_t captures cyclical or trend components of investment rates and is interacted with $D0_{jt}$ to allow for differences between unbanked and banked firms.

We further generalize the empirical model to allow for differences in the parameters of the Euler equation within the category of banked firms, to distinguish between single- and multiple-banked firms in the data. Equation [3] below includes interactions of the four observable explanatory variables with two dummy variables: DS, $S = 1, 2,$ where $D1$ ($D2$) equals 1 if the firm has only one (more than one) bank relation in a given year, and 0 otherwise:[41]

$$\left(\frac{I}{k}\right)_{jt} = \beta_1\left(\frac{I}{k}\right)_{jt-1} + \beta_2\left(\frac{I}{k}\right)^2_{jt-1} + \beta_3\left(\frac{q}{k}\right)_{jt-1} + \beta_4\left(\frac{y}{k}\right)_{jt-1} + d_t \qquad [3]$$
$$+ \sum_{s=1,2}\left[\delta_1^s\left(\frac{I}{k}\right)_{jt-1} \times DS_{jt} + \delta_2^s\left(\frac{I}{k}\right)^2_{jt-1} \times DS_{jt} + \delta_3^s\left(\frac{q}{k}\right)_{jt-1} \times DS_{jt} + \delta_4^s\left(\frac{y}{k}\right)_{jt-1} \times DS_{jt} + d_t{\times}DS_{jt}\right] + \eta_j + \varepsilon_{jt}.$$

where, as in [2], the coefficients β_l, $l = 1,..., 4$, are parameters for the reference group of unbanked firms, the coefficients $\delta^1{}_l$, $l = 1,..., 4$, quantify differences in the corresponding β_j between single-bank relation firms and unbanked firms, and the $\delta^2{}_l$, $l = 1,..., 4,$ quantify differences in the corresponding β_l between multiple-banked and unbanked firms. Possible heterogeneity in the intercept term d_t is also taken into account by interacting it with $D1$ and $D2$.

The potential endogeneity of the number of banking relationships, and consequently, of the interactions of the dummy variables across groups of firms (between banked and unbanked or among unbanked, single banked and multiple banked firms) in Equations [2] and [3], is treated through the

[40] This model corresponds to the first column of Table 15.
[41] This model corresponds to the second column of Table 15.

inclusion of the third lag of this variable in the list of instruments.[42] *A priori*, lagged decisions on the number of bank relations should be unrelated to current funding needs. For each specification we report the standard over-identifying restrictions Hansen-J test (robust to heteroskedasticity and autocorrelation) to test the validity of the list of instruments used.[43]

4.1.2. Investment Euler equations GMM results

Table 13 contains the parameter estimates for Equation [1] and their p-values (in brackets) for the whole sample. While the coefficients associated with the first lags of relative investment and squared relative investment $\hat{\beta}_1$ and $\hat{\beta}_2$ satisfy the restrictions imposed by the theory under the null of no financial constraints (they are significant; have the expected signs and the 95% confidence intervals for both parameters include a range of values higher than 1 in absolute value), the estimate of the cash flow coefficient, $\hat{\beta}_3$, is significantly *positive* (as opposed to negative as predicted by theory) for any conventional significance level. This 'excess sensitivity' of investment to cash flow is often interpreted as preliminary evidence that there are financially constrained firms in the sample (for which the standard investment Euler equation does not hold). The coefficient $\hat{\beta}_4$ on $\frac{y}{k}$ is positive and significant, consistent with absence of perfect competition in the output market.

The low p-value associated with the Hansen-J over-identifying restrictions test rejects the null that the orthogonality conditions are valid, indicating model mis-specification (Hansen, 1982). The null of absence of autocorrelation in the errors of the first-differenced equation cannot be rejected for orders greater than 2. This suggests an MA(l) structure of the error term in the levels equation and thus justifies the use of lags dated *t-3* and longer to instrument the endogenous variables in the first-differenced equation.

The results for the parameter estimates and test-statistics from a robustness check using the Blundell-Bond estimator reported in Appendix 4 (Table A4.1) are qualitatively similar to those obtained with the standard first-differenced Arellano and Bond GMM estimator. This suggests that potential problems with the Arellano and Bond estimator due to the presence of highly persistent series are not important in this model.

[42] Notice that under the presumption that the number of banking relationships is simultaneously decided with the investment rate by the firm, the inclusion of appropriate lags of the investment rates as instruments for the first differences of lagged and squared lagged investment rates in the right-hand side is also instrumenting the interaction terms differentiating across groups of firms with different number of bank relations.

[43] The list of instruments consists of third and fourth lags of *I/k and* the other endogenous or predetermined variables *q/k* and *y/k* and the third lag of the number of banking relations. Older lags have not been included in the instruments matrix for computational reasons.

We hypothesize that the presence and number of banking relationships affects the ability of the firm to obtain external funds to finance its investment projects. Consequently, we expect that distinguishing across categories of firms defined by their number of banking relations should be a useful criterion to elicit the unobservable categories of financially constrained and financially unconstrained firms in the whole sample. We assume that the decision whether to have or not have a banking relationship and the decision on the number of those relationships at time *t* are determined endogenously with the level of investment in period *t*.

Table 14 shows the Arellano and Bond GMM estimates[44] of the investment Euler equation for three alternative specifications. Columns (1) and (2) provide coefficient estimates for the specifications in Equations [2] and [3], respectively, which differentiate between the parameters of the Euler equation for unbanked and banked firms, in the former case, and among those for unbanked, single-banked and multiple-banked firms, in the latter case.[45]

The results in column (1) of Table 14 indicate that the coefficients for the reference category of firms in Equation [2], i.e., unbanked firms (*D0*=0), satisfy the restrictions imposed by the Bond and Meghir model. The 95% confidence intervals for the parameters associated with $(I/k)_{jt-1}$ and $(I/k)^2_{jt-1}$ include a range of values higher than 1 and the coefficient associated with the cash flow term, *q/k,* is negative and significant with a p-value of 2.7%, in line with the null of absence of financial constraints. The coefficient associated with the output term is positive and significant, with a p-value of 5.3%, consistent with imperfect competition in the output market.

The differences in the parameters of the Euler equation for banked firms (*D0*=1) relative to unbanked firms are significant for all coefficients, except for that on the output term, which suggests different models behind firms' investment decisions for the two categories. In particular, in contrast to the significantly negative coefficient for unbanked firms, the coefficient on the cash-flow rate for banked firms is positive (0.07, the result of the sum of the coefficient, β_3 for the reference group, -0.19, and the differential, $\delta_3 = 0.26$) and significant (the p-value associated to the test-statistic for the null $H_0 : \beta_3 + \delta_3 = 0$ is close to zero). This is traditionally interpreted as suggestive of the existence of financial constraints affecting firms in the banked category. The autocorrelation tests validate the use of third and further lags of the investment rate as instruments. Also, the Hansen-J test cannot reject the null that the instruments' matrix used is valid.

[44] The instruments matrix uses the 3rd and 4th lags of *I/k*, *q/k* and *y/k*. The potential endogeneity of the bank relations category to which each firm is assigned to has been taken into account by, additionally, considering the orthogonality conditions of the first -differenced error term with respect to the 3rd lag of the firm's number of banking relations.
[45] Each year firms are assigned to a group depending on the number of bank relations maintained by the firm in that particular year.

Column (2) of Table 14 provides the results from the GMM estimation of Equation [3], which estimates in a single regression the Euler equation for the reference group of unbanked firms, as well as the differences in the parameters for single-banked and multiple-banked firms relative to unbanked firms. The null that the moment conditions are valid cannot be rejected for any standard significance level (the Hansen-J test p-value is 74%). This, together with the autocorrelation tests analogous to those for column (1), globally validates the set of instruments for endogenous and predetermined variables. Inference from this specification allows us to conclude that differences in the parameters of the Euler equations for single-banked and multiple-banked firms relative to unbanked firms are concentrated in the coefficient on the cash flow term, as the null hypothesis that $\delta_l^s = 0$, for l=1 ,2 and 4; s=1, 2 cannot be rejected for any standard significance level (p-value is 28.3).

Given the results from specification (2), in column (3) of Table 14 we report GMM estimates for a restricted version of Equation [3] that imposes the null of no differences in the parameters of the Euler equations for single-banked and multiple banked firms relative to unbanked firms, except for the cash flow term coefficients. The results for unbanked firms, used as the reference group (D1=0 and D2=0), are similar to those discussed for column (1). The coefficients measuring the difference in parameters between the Euler equations for single- and multiple-banked firms relative to unbanked firms (δ_3^1 and δ_3^2), are compatible with the existence of further differences *within* the category of banked firms. Specifically, for single-banked firms the cash flow coefficient is positive (0.20, equal to the sum of -0.11 and 0.31) and significant (p-value of 0.4%) while for multiple-banked firms it is not significantly different from zero (p-value of 92%). Thus, unlike the result for unbanked firms, we do not find evidence that the cash flow coefficient on multiple-banked firms is negative as would be predicted under the null of no financial constraints. These results qualify the conclusion that banked firms are constrained while unbanked firms are not and show that the 'intensity' of financial constraints appears to vary with the number of banking relations. The autocorrelation and the Hansen-J tests confirm the validity of the list of instruments.[46]

Differences across firms with different trajectories in and out of the unbanked status

In Table 15, we group firms in four categories[47]: continuing unbanked firms,[48] continuing single-banked firms, continuing multiple-banked firms and rest of the firms.[49] We are interested in

[46] To check robustness we have also estimated the specifications reported in Table 14 using the Blundell and Bond GMM estimator. Results are provided in Appendix 4 (Table A4.2). Qualitative conclusions that can be drawn from these estimations are unchanged with respect to those discussed for Table 14. Similarities are especially remarkable for the restricted version of specification provided in columns (3) of Tables 14 and A4.2. However, for the Blundell and Bond GMM estimations, the p-values for the Hansen-J overidentifying restrictions test rejects the null that the orthogonality conditions as a whole are valid. Since, on the contrary, the Hansen-J tests do not reject the null that the orthogonality conditions for the Arellano and Bond estimator are valid, we interpret that the additional orthogonality conditions (assumptions) implicit in the Blundell and Bond estimator are likely invalid.

[47] This categorization coincides with that considered in Section 4.2.

examining differences in the parameters of the Euler equation across these groups of firms.[50] Table 15 reports the (Arellano-Bond) GMM estimates of these parameters. The autocorrelation and Hansen-J tests validate the specification and the list of instruments.

The most remarkable finding is, once again, the different sign of the coefficient associated to the cash flow term across the different firms. For continuing unbanked firms this coefficient is negative and significant. In spite of the highly imprecise estimates of the rest of the parameters,[51] this can be considered as evidence on the absence of financially constrained firms within this category. On the contrary, for the category of rest of the firms, all theoretical restrictions on the parameters of the Euler equation are satisfied except that on the sign of the coefficient on the cash flow term, which is positive and significant suggesting the presence of financially constrained firms in that group.

Finally, for the groups of continuing-single and continuing-multiple banked firms the coefficient associated to the cash flow term is, as the rest of the coefficients, not significant. This suggests that the Euler equations framework is not appropriate to describe the investment behaviour of firms in these categories. One possibility is that Euler equations are not appropriate due to the presence of financially constrained firms in these groups. Though conclusions for the continuing categories are weaker than those discussed for the non-continuing categories, they are compatible with our previous findings. In Section 4.2 we provide further evidence based on structural estimates for this categorization of firms that complement and reinforce results presented here.

Appendix 5 provides several robustness checks using the categories of continuing status firms using an alternative list of instruments, and using the Blundell-Bond approach. Our baseline results reported in this section remain robust. [52]

4.2. Structural Estimation

[48] We define as continuing unbanked firms those that are unbanked for most of the years in the sample (six or more years). Although we do not report the results in Table 15, we also did estimation runs where we defined continuing unbanked firms more strictly as never banked in the studied period. The results are similar but the significance of the estimates deteriorates as the number of observations goes down.

[49] The category of the rest of the firms includes both “switching” firms and “other continuing banked” firms (for definitions see Section 3).

[50] We assume that coefficients associated to I/k, $(I/k)^2$, q/k and y/k are heterogeneous across categories. We restrict the intercept term d_t to be equal for all the categories. More generally, heterogeneity in this term can also be taken into account. We impose the restriction of homogeneity since if this assumption is relaxed the estimates are less precise.

[51] The coefficients on I/k and $(I/k)^2$ corresponding to continuing unbanked firms are highly non-significant. However, the null hypotheses that these coefficients are equal to those for the category of rest of the firms, which satisfy theoretical restrictions, cannot be rejected (p-values 0.66 and 0.36, respectively).

[52] For robustness purposes we also explored other strategies to treat the potential endogeneity of the continuing categories. Specifically, we estimated the Euler equation for the four groups of firms including the Herfindahl concentration index of bank loans in the province of the firm, in addition to the 3rd and 4th lags of I/k, q/k and y/k. The main conclusions do not vary much with respect to those discussed in this section. Additionally, we estimated the Euler equation excluding the “rest of the firms” category. In this case, we postulate that since only firms that maintain a constant number of banking relations over the entire period are taken into account, the status of the firm can be considered exogenous. While the parameter estimates are similar and our main conclusion on the different sign of the cash flow coefficient across groups of firm is maintained, most parameter estimates become non-significant since the number of observations is greatly reduced.

4.2.1. Empirical Method

In this section we estimate the four dynamic models of firm credit access and finance developed in Section 2: autarky (A), non-contingent borrowing (B), moral hazard (MH), and full information / complete markets (FI) using structural maximum likelihood techniques as in Karaivanov and Townsend (2010).

We write down a likelihood function that measures the goodness-of-fit between the data and each of the four alternative models of financial markets. We then use the maximized likelihood value for each model (at the maximum likelihood estimated parameters) and perform a formal test (Vuong, 1989) about whether we can statistically distinguish between each pair of models relative to the data. We thus approach the data as agnostic about which theoretical model fits best and let the data determine this. The results of the Vuong test, a sort of 'horse race' among the four models, inform us which model(s) fits the data best and also which models can be rejected as likely to have generated the data.

Specifically, we have the panel data $\{\hat{k}_{jt}, \hat{I}_{jt}, \hat{q}_{jt}\}$ for j=1,...,n and t=0,...,T, where subscripts j and t denote firm and time, respectively; k is capital, I is investment, and q is cash flow. We use the first-period capital data $\{\hat{k}_{j0}\}$ to compute (via histogram function) the initial distribution of capital over the capital grid K. This distribution, denoted $H_0(\hat{k}) \equiv \{h_0(k_1),...,h_0(k_{\#K})\}$ is needed to initialize the each of the four regimes since k is state variable in each of them. For the initial values of the other (unobservable) state variables, b (debt) and w (promised utility) in the B, MH and FI regimes, we assume that they are drawn independently from k from the Normal distributions[53] N(μ_b, γ_b) and N(μ_w, γ_w) respectively and discretized over the debt/savings and promises grids B and W. The distribution parameters μ_b, γ_b or μ_w, γ_w are estimated together with the structural parameters σ, θ and ρ.

Next, we take the data, $\{\hat{x}_j\}_{j=1}^n$, i.e., the triple $\{\hat{k}_{jt}, \hat{I}_{jt}, \hat{q}_{jt}\}$ for a given time period t, and form the likelihood function implied by each regime. The basic idea is as follows. The solution to the dynamic linear program for each regime for a given vector of structural (σ, θ, ρ) and unobserved state distribution (μ_b, γ_b or μ_w, γ_w) parameters and for a given initial distribution for the observed state from the data $H_0(\hat{k})$ implies a joint distribution over the possible values k, I and q can take. Suppose the model grids imply a total of $G = \#K \times \#I \times \#Q$ distribution 'cells' (bins). Call these cells X_l , l = $1,...,G$. The MLE method minimizes the "distance" between this joint distribution for each model and its counterpart in the data. In other words, we evaluate and compare the fit with the data of each

[53] This particular distributional assumption is not essential for our method. Alternative specifications over the initial joint distribution of (k,b) or (k,w) can be used at the cost of additional parameters to be estimated and computational time.

model based on the predicted firm investment behavior (I) jointly with cash flow (q) conditional on firm's asset (k).

Formally, let the model $m \in \{A, B, MH, FI\}$ which we estimate generate the joint probability distribution $f^m(x \mid \phi, H_0(\hat{k}))$ over the vector of variables x (i.e., k, I, q), given a vector of parameters (the structural parameters σ, θ, ρ and the distributional parameters μ, γ for the unobserved state variables debt, b or promised utility, w), and given the initial distribution over the observable state $H_0(\hat{k})$. We also allow for normally distributed additive measurement error in each variable with mean zero and standard deviation parameterized by γ_{me}. This parameter is also included in the overall vector of estimated parameters ϕ. The addition of measurement error results in the joint distribution $\tilde{f}^m(x \mid \phi, H_0(\hat{k}))$.[54] Altogether this implies that we estimate the following six parameters, $\phi \equiv \{\sigma, \theta, \rho, \gamma_{me}, \mu_{b/w}, \gamma_{b/w}\}$.[55]

Since our computational methods require discretization (the use of grids) for all variables, the joint distribution over k, I and q our dynamic models generate is discrete. Moreover, this distribution representing the solution of the linear program is non-trivial (due to the uncertainty in q and possible convexification lotteries) and, because of the dynamic structure and complexity of our models, it has no analytic form. The joint distribution over $x \equiv (k, I, q)$ obtained after adding measurement error also has no analytic form due to the unknown underlying joint distribution $f^m(x \mid \phi, H_0(\hat{k}))$ and the multi-dimensionality of the distribution. This implies that the model likelihood at the (theoretically continuous) data is analytically intractable and has to be simulated. Therefore, for computational reasons, to construct the model likelihood at the data we need to also discretize the joint distribution of k, I, q in the data over the same grid cells $\{X_l\}$ used to compute the models. Call this discretized distribution of k, I, q in the data $\{\hat{x}_j\}_{j=1}^n$. For example, in our baseline estimation runs we use grids of five points for each k, I and q. These grids imply 125 mutually exclusive cells/bins, X_l (or, equivalently, a 5-by-5-by-5 three-dimensional grid over k, I and q), the observed frequencies over which represent the joint distribution of the data $\hat{x}_j = \{\hat{k}_{jt}, \hat{I}_{jt}, \hat{q}_{jt}\}$.

The likelihood of the observed data for model m, assuming the data are i.i.d. over firms, is then:

[54] The measurement error is added independently for each k, I and q using the Normal pdf and a histogram function to determine the transformation of $f^m(x \mid \phi, H_0(\hat{k}))$ into $\tilde{f}^m(x \mid \phi, H_0(\hat{k}))$ by essentially computing how much of the probability mass 'spills' into nearby cells due to measurement error.

[55] Naturally, in the autarky regime only the first four parameters are estimated. Also, in the baseline runs with risk-neutral firms we set σ=0 and do not estimate it.

$$L(\phi) = \prod_{j=1}^{n} \widetilde{f}^{m}(\hat{x}_j \mid \phi, H_0(\hat{k}))$$

Taking logs, we can rewrite this as:

$$\Lambda(\phi) = \sum_{j=1}^{n} \ln \widetilde{f}^{m}(\hat{x}_j \mid \phi, H_0(\hat{k})) =$$
$$= \sum_{j=1}^{n} \sum_{l=1}^{G} 1(\hat{x}_j \in X_l) \ln prob^{m}(x \in X_l \mid \phi, H_0(\hat{k})) \qquad \text{(LL)}$$

where $prob^{m}(x \in X_l \mid \phi, H_0(\hat{k}))$ denotes the probability implied by model *m* that the vector of variables *x* belongs to each given cell X_l, given the initial observable state distribution $H_0(\hat{k})$ and given parameters ϕ and where 1(.) is the indicator function. We maximize the log-likelihood (LL) by choice of ϕ for each of the four dynamic financial regimes (A, B, MH, and FI) from Section 2. Standard errors are computed via bootstrap, repeatedly drawing with replacement from the data up to the original sample size.

After we estimate and obtain the maximized likelihood (LL) for each model, we follow Vuong (1989) and compute an asymptotic test statistic that we use to formally distinguish (in bilateral comparisons) across the alternative theoretic financial regimes. The Vuong test does not require that either of the compared models be correctly specified. If the two compared models are non-nested (such as ours, see Vuong (1989) for formal definition), the Vuong test-statistic is normally distributed under the null that the two models are equally close to the data. If the null is rejected (the Vuong Z-statistic is large enough in absolute value), we say that the higher likelihood model is closer to the data than the other.

4.2.2. Baseline parameters and functional forms

For utility we use the functional form $u(c,z) = \frac{c^{1-\sigma}}{1-\sigma} - z^{\theta}$, where $\sigma \geq 0$ is the risk aversion parameter (with σ=0 corresponding to risk neutrality) and θ is the curvature of the disutility of effort.

On the production side, suppose net profits $q(k, z, \theta^{q})$ can take on a finite number of realizations, q_i, $i = L, \ldots, \#Q$ and disregard the shock θ^{q}. Then for q_i at its lowest value q_L, let

$$\text{Prob}\,(q = q_{\text{L}} \mid z, k) = 1 - \left(\frac{k^{\rho} + z^{\rho}}{2}\right)^{\frac{1}{\rho}},$$

while for any other q_i, $i \neq L$, let,

$$\text{Prob}\,(q = q_{\text{i}} \mid z, k) = \frac{1}{\#Q - 1}\left(\frac{k^{\rho} + z^{\rho}}{2}\right)^{\frac{1}{\rho}}.$$

There are constant returns to scale in k and z. We assume equal weights for capital and labor in the CES specification. For simplicity, assume $q_L > 0$ so that some borrowing is always possible, with the loan being repaid out of earnings even when the firm is in the worst q_L state forever. For example, with $\#Q=2$ (so that either $q = q_L$ or $q = q_H > q_L$) expected profits are CES in capital and effort, with ρ the elasticity of substitution: $\rho < 0$ means effort z and capital k are substitutes, $\rho = 0$ is the Cobb-Douglas special case, and $\rho > 0$ implies inputs are complements.

The above functional forms are not necessary for our computational and estimation methods. Our linear programming approach is extremely general and can accommodate any functional form (including non-convex) for preferences and technology. The main reason we picked the particular forms we use here is their parsimonious use of parameters (key for computation speed).

To compute and estimate the four models with our yearly data we calibrate β=.95, R=1/.95. We also calibrate δ=.16 which equals the median tangible asset depreciation obtained from the Spanish data.[56] We convert the data into model units (normalize) by dividing all currency values for *k, I* and *q* by the 90-th percentile of the distribution of tangible fixed assets in the data (referred to as "capital", *k*). We use the grids *K, I* and *Q* to form the log-likelihood criterion (LL), and perform the MLE estimation. In our baseline runs, we use the five-point grid {.015, .05, .12, .3, 1} for capital, *k*.[57] For cash flow *q* we use the five-point grid {.007, .02, .05, .11, .42} and for investment *I* we use the five-point grid {.0004, .006, .017, .05, .22}. The grid bounds were chosen to reflect the actual data relative magnitudes while the grid spacing matches the data quintiles to try ensure a sufficiently large number of observations in each *K*x*Q*x*I* grid cell. To be consistent with the Bond-Meghir Euler equation approach our baseline estimation runs (Table 16) focus on the case of risk neutral firms (that is, we set σ=0 and do not estimate it), but in the robustness checks (Table 18) we also report results allowing for risk-averse firms.

4.2.3. Results

Table 16 contains our baseline model comparison results obtained using the Vuong test. Each column contains the Vuong test statistics for our baseline specification comparing each financial regime to the other three (six possible pairs in total). A statistically significant test statistic (in the Table we report 1%, 5%, and 10% significance levels) indicates that we can reject the null hypothesis of both models being equally close to the data, i.e., the model listed in the parentheses fits the data better than the alternative model to which it is compared.

[56] We estimated the empirical distribution function of firm-level depreciation rates (measured by firm's accountant depreciation expressed as percentage of firm's tangible fixed assets) for 1997-2004. For each year, we obtained the median depreciation rate across firms and we set the value of δ equal to the average of these median depreciation rates.

[57] We use a standard histogram function based on distance to the closest grid point (Matlab's command *hist*). Our methods can handle much finer grids at the cost of computation time.

The first section of Table 16 uses data from 1997, the initial year of our panel. The findings indicate that in the whole sample, under risk-neutrality, the compared models rank in terms of best to worst as: borrowing (B), moral hazard (MH) tied with full information (FI) and autarky (A). That is, the non-contingent debt borrowing model comes closest to the data in terms of its likelihood based on the joint distribution of (k, I, q) data in 1997.

We perform the same exercise when the data are stratified by firm type according to the number of bank relations. The results (also reported in Table 16) reveal patterns that confirm and reinforce our findings from the Euler equation GMM analysis in Section 4.1. and Table 15. Namely, we compare the fitness of the financial regimes with the data across four categories of firms, defined according to the number of relations with banks over the entire period 1997-2004. Specifically, we explore differences among the categories of "continuing unbanked" firms,[58] "continuing single-banked" firms and "continuing multiple-banked" firms. To these we add a fourth residual category that includes the remainder of the firms in the sample (this group includes "switching firms" and "other continuing banked firms"), referred to in Table 17 as "others". This categorization of firms allows taking into account the endogeneity of the number of bank relations since for these continuing-status firms, their dynamic programming problem implies that they would behave approximately as if they were to stay in that category forever, especially if at the end of the sample they are not near making a transition.

Our results show that for continuously unbanked firms the moral hazard (MH) model fits the 1997 capital, investment and cash flow data best (the first-best unconstrained regime, FI comes in at a statistically significant second place). In stark contrast, continuing single-banked firms, continuing multiple-banked and firms in the residual category are each revealed to be the more financially constrained, with the borrowing, non-contingent debt model coming out on top in terms of fit with the data, relative to the moral hazard, full information and autarky regimes. These results confirm the evidence of financial constraints affecting these three categories relying on the Euler equations framework, contrarily to firms in the continuing unbanked category (Table 15 in Section 4.1.2). We obtain very similar results for the four firm types by number of banking relationships when we use 2004 (k,I,q) data (Section 2 in Table 16). One minor difference is that now the full information, perfect credit markets regime comes very close to the moral hazard constrained regime for the unbanked firms (distinguishable only at the 10% level).[59]

In Table 17 we report the MLE estimates and bootstrap standard errors for each of the four regimes and each of the four data stratifications by banking status. Though the borrowing regime fits the data

[58] Here we use the less restrictive definition of continuing unbanked firms that includes firms that are unbanked six or more years during the period 1997-2004 as in Table 15 earlier on.

[59] Also, in the whole sample, moral hazard is no longer tied with the FI regime.

best for the categories of single-banked, multiple-banked and the rest of the firms, the estimated parameter values vary, sometimes significantly, across the stratifications which is the way the structural model accommodates the diversity in the data across the firms. The same is true within stratifications looking across the different models.

4.2.4. Robustness

To check robustness, we also estimate and carry out the Vuong model comparison tests when we allow for quadratic adjustment costs, risk aversion, and alternative variable grids (Table 18). We first explicitly include quadratic adjustment costs to investment in each of the four financial regimes using the specification for $g(k, I; \theta^I)$ from Section 2. The cost g is subtracted from dividends c in the optimization problem. We estimate the adjustment cost parameter b (it is included in the vector ϕ) and, for simplicity, set ς, θ^I and b_i to zero. Our baseline results (without adjustment costs) reported in Table 16 can be interpreted as fixing $b=0$. Also, the FI regime with adjustment costs can be interpreted as a "pure adjustment costs" regime, with no other financial constraints, as in the Bond and Meghir (1994) model.

Including explicit adjustment costs (weakly) improves the fit to the data of all four financial regimes and all four firm stratifications by banking status. The improvement in likelihood is largest for the full information (FI) regime and smallest for the B and A regimes where in a few cases the likelihood maximization chose to set the parameter b to zero corresponding to no adjustment costs. Still, the improvement in the likelihood of the MH and FI regimes with adjustment costs for single-banked, multiple-banked and other firms is insufficient to displace the B regime as the best-fitting one. As a result, the best fitting regimes when such exogenous capital adjustment costs are allowed does not change compared to the no adjustment cost benchmark – compare Table 16, Section 1 and Table 18, Section 1. The only minor exception is the tie between MH and FI in the estimation runs with adjustment costs for continuously unbanked firms. This means that we cannot reject the hypothesis that, allowing for adjustment costs, continuing unbaked firms are completely unconstrained. This finding is consistent with our GMM results.

Section 2 of Table 18 shows that our baseline results for the group of continuously unbanked firms are not dependent on the assumption of risk neutrality, which is implicit but key in the Euler equations approach and most of the literature. Specifically, when we allow for risk aversion ($\sigma > 0$) and include the risk aversion coefficient σ in the list of estimated parameters, we find that the borrowing model (B) is, once again, best-fitting for the "single-banked", "multiple-banked" and "others" firm groups. As in the baseline runs with risk neutral firms, the moral hazard regime fits the data best for continuously unbanked firms.

Finally, we also did runs with a set of different, coarser variable grids (three points each for k, i, q and b). The Vuong test statistics are presented in Section 3 of Table 18. The coarser grids inhibit somewhat our ability to distinguish among the financial regimes relative to the data (we observe more model ties relative to the baseline). Still, the results remain consistent with the baseline runs. The moral hazard regime achieves the highest likelihood (although here it is statistically tied with FI) for unbanked firms, while the non-contingent debt (B) regime still achieves the highest likelihood for single- and multiple-banked firms.

5. How do unbanked firms finance their investment projects?

Since unbanked firms in our sample do invest and we find evidence compatible with the absence of financial constraints for firms in this category, a pertinent question is how do they obtain the necessary funds, in addition to internal resources, to finance their business? In this section, we provide detailed evidence on this question using two complementary approaches. First, we use SABI-INFORMA to itemize debt items from the firms' public financial statements and draw some conclusions from the aggregated numbers. Second, we perform detailed "case studies" of formal and informal relationships with other firms for some of the large continuously unbanked firms in our sample. To do this we explore the information provided by SABI-INFORMA on members of the firms' boards and the firms' shareholders and subsidiaries.[60]

5.1. Aggregated Debt

Information about the itemization of debt as reported in financial statements is non-existent for most firms because by law only audited firms must report such detailed information. Still, given our findings on unbanked firms above, it seems important to do our best to find out as much as possible about the structure of their debt. Taking as an example 1997, we have data on the itemization of debt for 2,468 firms, out of a total of 236,301 firms.[61] Within these firms, a further 10% (248 firms) report no borrowing from banks.[62] Restricting attention to this sub-sample of unbanked firms, we aggregate their debt item by item to learn about the sources of funds used by unbanked firms to finance investment projects. The results are presented in Table 19.

[60] This information is static in SABI-INFORMA in the sense that it is available only for the last report by the firm. The results presented in this section are based on the October 2009 update of the database.
[61] Structural characteristics are stable over time.
[62] Here we refer to the information on bank loans reported in firms' public financial statements.

In aggregate, around 72% of unbanked firms' debt is provided by other *associated and affiliated companies*. The next important source of funds is *trade credit,*[63] followed by *government funding* (for example, delay in tax payments), which represent 10.5% and 3.5% of aggregate debt, respectively. In contrast, the two most important sources of funds for banked firms are *bank loans* and *trade credit* representing approximately 34% and 29% of total debt, respectively. 'Other firms in the group' as source of debt comes only third for banked firms representing 17.1% of aggregate debt, much less than the 72% for the unbanked group. Admittedly, these numbers come from a relatively small number of firms but they represent our most accurate firm debt decomposition data.

5.2. Case Studies

We also thoroughly studied 18 large continuing unbanked firms randomly drawn from the list of 25 largest continuing unbanked firms according to total assets in 2004. The results are presented in Appendix 6. For each firm, we used the SABI-INFORMA database that contains all available financial and non-financial information reported by that firm. In particular, the firm's report lists all of its shareholders and subsidiaries. When a shareholder or subsidiary is also contained in the SABI-INFORMA database, the two records were linked. Thus, by manually combining firm reports we managed to recover networks of formally related firms. In addition, we discovered other relations across firms in our data, taking into account connections across firms' boards (e.g., whether the president and/or chief executive are the same, or whether they are family members) and other "coincidences" such as, common address and telephone number (suggesting a more "informal" group structure). For firms in any such (formal or informal) structure, we complemented the information drawn from SABI-INFORMA with that from CIR to check whether or not these related firms are banked.

The above procedure, while time-consuming, allows us to conclude that persistently unbanked firms in our sample belong, to a significant extent, to formal complex structures of firms within which firms are directly or indirectly linked by shareholder-subsidiary ties. Another frequent pattern is for linked firms to be owned and/or managed by the same person or by members of the same family. These familial structures of firms can be *pyramidal* (in which predominant ties across the firms are of the type shareholders-subsidiaries) or what we call *informal*. The latter case occurs if board members are the same or belong to the same family or if the firms report the same address or telephone/fax number. In general, these family firms are well capitalized and do not need banks' funding to develop their business plans, although occasionally some firms within their structure get loans from banks. Finally, looking at the object of business of these firms, we see significant diversity not only across firms but

[63] Carbó, Rodriguez and Udell (2008) report evidence on the use of bank loans by unconstrained firms to finance trade credit provided to other firms.

also across members of the same group. Appendix 6 provides a detailed description of our findings for one illustrative case.

6. Summary and conclusions

We use data from the Spanish Mercantile Registry and from the Bank of Spain Credit Registry (CIR) to classify firms according to whether they have loans from a formal Spanish credit institution, from several such institutions, or from none. We restrict attention to firms that are at least ten years old and find, nevertheless, that a significant number are unbanked. We find evidence that these firms' debt is non-trivial, but comes largely from borrowing from other firms which are members of the same group. In the case of persistently unbanked firms, we document that many belong to family groups which facilitate funding through channels alternative to formal bank credit. We then examine the existence and magnitude of financial constraints for each firm type (unbanked, single-banked or multiple-banked) using two related but distinct approaches based on a common theoretical model.

First, we use a standard Euler equation approach to investment. We find that single-banked firms are most likely to exhibit excess investment cash flow sensitivity. Unbanked firms are not. Our findings are consistent with the literature arguing that access to bond markets comes only for larger firms. However, in Spain higher-level finance beyond bank loans is not obtained issuing bonds but from the Spanish analogue to *chaebols* and *zaibatsu* (i.e., large and diversified groups of firms). We further find weak evidence for differences in cash flow sensitivity between single-banked and multiple-banked firms suggesting that the intensity of financial constraints affecting investment decisions is likely higher for single-banked firms.

In our second, complementary empirical approach, we estimate four dynamic models of non-nested financial regimes ranging from complete markets to financial autarky for firms differing in their number of continuing banking relations using simulated maximum likelihood. Our results from this structural approach are very much in line with the results from the Euler equation GMM. Using Vuong's (1989) model comparison test, we find that continuously unbanked firms have investment behavior that fits best a moral hazard-constrained credit model (statistically tied with complete markets when adjustment costs are allowed), while continuing single-banked and continuing multiple-banked firms operate in a financial environment that best fits a non-contingent debt regime.

Our contribution in this paper is two-fold. First, to the best of our knowledge, the combination of empirical approaches we use, Euler equations regressions and structural MLE estimation, is unique. Second, and more importantly, the two complementary approaches tell a consistent story of financial market imperfections faced by Spanish unlisted non-financial firms: single-banked and multiple-banked firms are more severely financial constrained than older unbanked firms.

References

Alonso-Borrego, C. (1994): "Estimating dynamic investment models with financial constraints", working paper nº 9418, CEMFI.

Arellano, M. and S. Bond (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *Review of Economic Studies*, 58, 277-297.

Banco de España (2006): *Memoria de la Central de Información de Riesgos*, 2005.

Berger, A. N. and G. F. Udell (1995): "Relationship lending and lines of credit in small firm finance", *Journal of Business*, 68, pp. 351-382.

Bertrand, M., P. Mehta and S. Mullainathan (2002): "Ferreting out Tunnelling: An application to Indian business groups", *Quarterly Journal of Economic*s, 117 (1), pp.121-148

Bertrand, M and S. Mullainathan (2003): "Pyramids", *Journal of the European Economic Association*, 1 (2-3), pp. 478-483.

Bertrand, M. and Schoar, A. (2006): "The Role of Family in Family Firms", *Journal of Economic Perspectives*, 20, 2, pp. 73-96.

Blundel, R. and S. Bond (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87, pp.115-143.

Bond, P. (2004): "Bank and non-bank financial intermediation," *Journal of Finance,* 59, 6, pp. 2489-2529.

Bond, S. (2002): "Dynamic panel data models: A guide to micro data methods and practice," *Portuguese Economic Journal*, 1, 141-162.

Bond, S. and C. Meghir (1994): "Dynamic investment models and the firms' financial policy"; *Review of Economic Studies*, 61, 2, 197-222.

Bond, S., J.A. Elston, J. Mairesse, and B. Mulkay (2003): "Financial factors and investment in Belgium, France, Germany and the United Kingdom: A comparison using company panel data, *Review of Economics and Statistics*, 85, 1, 153-165.

Boot, A.W.A. (2000): "Relationship banking: What do we know?", *Journal of Financial Intermediation*, 9, pp. 7-25.

Carbó, S., F. Rodríguez. and G.F. Udell (2008): "Bank lending, financing constraints and SME investment", WP 2008-04, Federal Reserve Bank of Chicago.

Chirinko, R.S. and J.A. Elston (2006), "Finance, control and profitability: the influence of German banks", *Journal of Economic Behaviour and Organization*, 59, pp. 69-88.

Cleary, S. (1999): "The relationship between firm investment and financial status", *Journal of Finance*, 54, pp. 673-692.

Cooper, R. and J. Ejarque (2003): “Financial Frictions and Investment: A Requiem in Q”, *Review of Economic Dynamics*, 6, pp. 710-728.

Cronqvist, H and M. Nilson (2003): “Agency costs and minority shareholder”, *Journal of Financial and Quantitative Analysis*, 38, 4, pp. 695-719.

Degryse, H. and Van Cayseele, P. (2000):”Relationship lending with a bank-based system: Evidence from European small business data”, *Journal of Financial Intermediation*, 9, pp. 90-109.

Diamond, D. W. (1984): “Financial intermediation and delegated monitoring”, *Review of Economic Studies*, 51, 393-414.

Elston, J.A. (1996): “Investment, liquidity constraints and bank relationships: evidence from German manufacturing firms”, Discussion Paper nº 1329, CEPR.

Elyasiani, E. and Goldberg, L.W. (2004): “Relationship lending: a survey of the literature”, *Journal of Economics and Business*, 56, pp. 315-330.

Estrada, A. and J. Vallés (1998): “Investment and Financial Structure in Spanish Manufacturing Firms”, *Investigaciones Económicas*, 22 (3), pp. 337-359.

Fazzari, F. M., Hubbard, R.G. and Petersen, B.C. (1988): “Financing Constraints and corporate investment”, *Brooking Papers on Economic Activity*, 1, 141-206.

Fazzari, F. M., Hubbard, R.G. and Petersen, B.C. (2000): “Investment-cash flow sensitivities are useful: a comment on Kaplan and Zingales”, *Quarterly Journal of Economics*, 115, 2, 695-705.

Fohlin, C. (1998): “Relationship banking, liquidity and investment in the German industrialization”, *Journal of Finance*, 53, 5, pp. 1737-1758.

Fuss, C. and P. Vermeulen (2006): “The response of firms’ investment and financing to adverse cash flow shocks: the role of bank relationships”, working paper nº 658, European Central Bank.

Gan, J. (2007): “The real effects of asset market bubbles: Loan- and firm-level evidence of a lending channel”, *Review of Financial Studies*, 20, 5, pp. 1941-1973.

García-Marco, T. and C. Ocaña (1999): “The effect of bank monitoring on the investment behaviour of Spanish firms”, *Journal of Banking and Finance*, 23, pp. 1579-1603.

Gomes, J., (2001): “Financing Investment”, *American Economic Review*, 90, 5, pp. 1263-1285.

Guiso, L. and R. Minetti (2007): “The structure of multiple credit relationships: Evidence from US firms”, Working Paper nº 2007/46, European University Institute.

Hansen, L.P. (1982): “Large sample properties of Generalised Method of Moments estimators”, Econometrica, 50, 4, pp. 1029-1054.

Harhoff, D. and T. Körting (1998): “Lending relationships in Germany: empirical evidence form survey data”, *Journal of Banking and Finance*, 22, pp. 1317-1353.

Hernando, I. and A. Tiorno (2002): “Financial constraints and investment in France and Spain: A comparison using firm-level data” in *French and Spain industrial corporations over the period 1991-1999: a comparative study based on their financial statements*, Central de Balances, Banco de España.

Hoshi, T, A. Kashyap and D. Scharfstein (1990): "The role of banks in reducing the costs of financial distress in Japan", *Journal of Financial Economics*, 27, pp.67-88.

Houston, J. F. and James, C. M. (2001): "Do relationship have limits? Bank relationships, financial constraints and investment", *Journal of Business*, 74, 3, pp. 347-374.

Hubbard, R.G. (1998): "Capital Market imperfections and investment", *Journal of Economic Literature*, 36, March, pp. 193-225.

Jaffee, D. and T. Russell (1976): "Imperfect information, uncertainty and credit rationing", *Quarterly Journal of Economics*, 91, pp. 651-666.

Jensen, M. and W. Meckling (1976): "Theory of the firm: managerial behaviour, agency costs, and ownership structure", *Journal of Financial Economics*, 1 pp. 305-360.

Jiménez, G. and J. Saurina (2004): "Collateral, type of lender and relationship lending as determinant of credit risk", *Journal of Banking and Finance*, 18 (9), pp. 2191-2212.

Jiménez, G.; V. Salas and J. Saurina (2006): "Determinants of collateral", *Journal of Financial Economics*, 81, pp. 255-281.

Jiménez, G.; J. A. Lopez and J. Saurina (2009): "Empirical analysis of corporate credit lines", Forthcoming, *Review of Financial Studies*.

Kaplan, S. N. and Zingales, L. (1997): "Do investment-cash flow sensitivities provide useful measures of financing constraints", *Quarterly Journal of Economics*, 112, 1, pp. 169-215.

Kaplan, S. N. and Zingales, L. (2000): "Investment-cash flow sensitivities are not valid measures of financing constraints", *Quarterly Journal of Economics*, 115, 2, pp. 707-712.

Karaivanov, A. and R. M. Townsend (2010): "Dynamic Financial Constraints: distinguishing mechanism design from exogenously incomplete market regimes", mimeo.

Khanna, T. and Y. Yafeh (2007): "Business Groups in emerging markets: paragons or parasites", Journal of Economic Literature, vol. XLV, june, pp. 331-372.

Maury, B. (2006):"Family ownership and firm performance: empirical evidence from Western European corporations", *Journal of Corporate Finance*, 12, pp. 321-341.

Myers, S. C. and Majluf, N. S. (1984): "Corporate financing and investment decisions when firms have information that investors do not have", *Journal of Financial Economics*, 13, pp. 187-221.

Morck, R.K. (2005): *A History of Corporate Governance around the world*, Randal K. Morck ed, The University of Chicago Press, Chicago and London.

Ongena, S. and D. Smith (2000): "What determines the number of bank relationships? Cross-country evidence", *Journal of Financial Intermediation*, 9, 1, 26-56.

Petersen, M. A., and Rajan, R. G. (1994): ''*The Benefits of Lending Relationships: Evidence from Small Business Data*'', *Journal of Finance*, 49, March, pp. 3–37.

Prescott, E.C. and R. Townsend (1984): "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard", *Econometrica*, 52, pp. 21-45.

Samphantharak, K. (2003): *Internal capital markets in business groups*, Thesis (Ph.D.) University of Chicago.

Schiantarelli, F. (1996): "Financial constraints and investment: Methodological issues and international evidence", *Oxford Review of Economic Policy*, 12, 2, pp. 70-89.

Schmid, L. (2009): "A Quantitative Dynamic Agency Model of Financing Constraints", *working paper*, Duke University.

Sraer, D. and Thesmar, D. (2004): "Performance and behaviour of family firms: Evidence from the French stock market", *Discussion Paper Series*, CEPR, London.

Stiglitz, J.E. and A. Weis (1981): "Credit rationing in markets with imperfect information", *American Economic Review*, 71, pp. 393-410.

Townsend, R. (1979): "Optimal contracts and competitive markets with costly state verification", *Journal of Economic Theory*, 21, pp. 265-293.

Townsend, R. and K. Ueda (2007): "Welfare gains from financial liberalization", *working paper* nº 07/154, IMF.

Vickery, J. (2005): "Banking relationship and the credit cycle. Evidence from the Asian financial crisis", mimeo.

Vuong, Q. (1989): "Likelihood ratio tests for model selection and non-nested hypotheses", *Econometrica*, 57, 2, 307-333.

Williamson, S. D. (1987): "Costly monitoring, optimal contracts and equilibrium credit rationing", *Quarterly Journal of Economics*, 102, pp. 135-145.

Table 1. Distribution of firms according to the number of years with available information.

Number of years	10	11	12
Frequency	7,571	12,299	16,715

Spanish non-financial firms: public and listed firms excluded.

Table 2. Distribution of firms according to the number of relations with banks

	1997	1998	1999	2000	2001	2002	2003	2004	average 1997-2004
unbanked	10.5	9.9	9.7	9.5	10.0	10.6	10.8	10.3	10.2
single banked	18.3	18.1	17.7	18.1	18.0	17.7	18.1	18.7	18.1
multiple banked	71.2	72.1	72.6	72.3	71.9	71.7	71.1	71.0	71.7
2 banks	19.2	18.8	19.0	19.6	19.1	18.7	18.5	18.9	19.0
3 banks	15.4	15.4	15.2	15.5	15.2	14.7	15.0	14.5	15.1
4 banks	10.8	10.7	10.8	10.8	10.6	10.7	10.7	10.8	10.7
5 banks	7.6	8.0	7.8	7.8	7.8	7.9	7.5	7.4	7.7
6 banks	5.3	5.3	5.5	5.5	5.5	5.4	5.2	5.3	5.4
7 banks	3.7	3.8	3.7	3.8	3.8	3.9	3.9	3.7	3.8
8 banks	2.6	2.7	2.8	2.7	2.9	2.9	2.8	2.9	2.8
9 banks	1.7	2.0	1.9	1.8	1.8	2.0	2.0	2.0	1.9
10 or more banks	4.9	5.3	5.8	4.9	5.2	5.5	5.5	5.5	5.3
Number of non financial firms[1]	36,585	36,585	36,585	36,585	36,585	36,585	36,585	36,585	36,585

[1]Spanish non-financial firms: public and listed firms excluded.

Table 3. Distribution of firms [1] according to the number of relations with banks, by age group (%).

Age group	Number of banking relations group	1997	1998	1999	2000	2001	2002	2003	2004	average 1997-2004
[10,20)	unbanked	9.8	9.2	9.4	9.4	10.9	10.9	10.8	11.1	10.2
[10,20)	single-banked firms	18.2	18.3	17.7	18.5	17.7	18.5	20.5	19.5	18.6
[10,20)	multiple-banked firms	72.0	72.5	72.9	72.1	71.4	70.6	68.8	69.4	71.2
[20,30)	unbanked	7.5	8.0	8.0	9.1	9.2	9.6	9.8	9.6	8.9
[20,30)	single-banked firms	15.1	14.6	15.5	15.4	16.4	16.1	16.5	17.8	15.9
[20,30)	multiple-banked firms	77.4	77.4	76.5	75.6	74.4	74.3	73.7	72.5	75.2
[30,∞)	unbanked	11.1	10.4	10.1	9.8	10.1	10.8	11.1	10.5	10.5
[30,∞)	single-banked firms	18.8	18.5	18.1	18.6	18.5	18.0	18.3	15.1	18.0
[30,∞)	multiple-banked firms	70.1	71.1	71.8	71.7	71.4	71.1	70.6	76.4	71.8
n.a	unbanked	11.1	10.4	10.1	9.8	10.1	10.8	11.1	10.5	10.5
n.a	single-banked firms	18.8	18.5	18.1	18.6	18.5	18.0	18.3	19.1	18.5
n.a	multiple-banked firms	70.1	71.1	71.8	71.7	71.4	71.1	70.6	70.4	71.0

[1]Spanish non-financial firms: public and listed firms excluded

Table 4. Transition matrices across unbanked firms, single-banked firms and multiple-banked firms (%).

One-year time horizon

		category in year t-1			
		single banked	multiple banked	unbanked	Total
category in year t	single banked	67.8	5.3	20.0	18.1
	multiple banked	21.3	93.6	7.5	71.8
	unbanked	10.9	1.1	72.5	10.1
	Total	100.0	100.0	100.0	100.0

Two-year time horizon

		category in year t-2			
		single banked	multiple banked	unbanked	Total
category in year t	single banked	56.2	7.5	25.2	18.1
	multiple banked	29.5	90.7	11.9	71.8
	unbanked	14.3	1.8	62.9	10.2
	Total	100.0	100.0	100.0	100.0

Five-year time horizon

		category in year t-5			
		single banked	multiple banked	unbanked	Total
category in year t	single banked	43.2	10.3	29.8	18.2
	multiple banked	38.9	86.4	20.5	71.2
	unbanked	17.9	3.3	49.7	10.6
	Total	100.0	100.0	100.0	100.0

Average percentages for 1997-2004.

Table 5. Distribution of firms[1] according to the number of years in which the firm is classified as unbanked during period 1997-2004.

	Number of years	Absolute frequency[1]	Relative frequency[2]	Relative frequency[3]
Continuing banked	0	28,050	76.7	100.0
Continuing single-banked	0	914	2.5	3.3
Continuing multiple-banked	0	19,083	52.2	68.0
Other continuing banked firms	0	8,053	22.0	28.7
Switching firms	1	2,440	6.7	31.9
	2	1,459	4.0	19.1
	3	1,096	3.0	14.3
	4	838	2.3	11.0
	5	709	1.9	9.3
	6	617	1.7	8.1
	7	480	1.3	6.3
	[1,7]	7,639	20.9	100.0
Continuing unbanked	8	896	2.4	
All firms		36,585	100.0	

[1]Spanish non-financial firms: public and listed firms excluded
[2]Relative to all the firms.
[3]Relative to the number of firms in the subtotal (continuing banked firms or switching firms).
Annual averages for 1997-2004.

Table 6. Distribution of 'switching' firms according to the number of changes in the banked/unbanked status over the studied period.

Number of changes	Absolute frequency	Relative frequency[2]	Relative frequency[3]
1	3,264	42.7	8.9
2	2,856	37.4	7.8
3	964	12.6	2.6
4	440	5.8	1.2
5	94	1.2	0.3
6	21	0.3	0.1
7	0	0.0	0.0
Switching firms	7,639	100.0	20.9
All firms[1]	36,585		100.0

[1]Spanish non-financial firms: public and listed firms excluded
[2] Relative to the subsample of switching firms.
[3] Relative to the entire sample of firms.

Table 7. Descriptive statistics for the distribution of age in unbanked, single-banked and multiple-banked firms (%).

	number of observations[1]	mean	standard deviation	min. value	p5	p25	p50	p75	p95	max. value
All firms	9,956	22.4	10.9	13.5	13.5	15.5	18.5	25.5	42.5	130.5
Single-banked firms	1,700	21.4	10.1	13.5	13.5	15.4	17.9	23.8	40.0	103.0
Multiple-banked firms	7,329	22.8	11.3	13.5	13.5	15.5	19.5	26.3	43.6	130.5
Unbanked	927	21.1	9.4	13.5	13.5	15.0	18.0	23.6	38.3	89.4

[1]Spanish non-financial firms: public and listed firms excluded
Annual averages for 1997-2004.

Table 8. Descriptive statistics for the distribution of size (total assets) for unbanked, single-banked and multiple-banked firms (thousands of Euros).

	number of observations[1]	mean	standard deviation	min. value	p5	p25	p50	p75	p95	max. value
All firms	36,585	7,447	61,195	11	235	659	1,462	3,652	20,718	4,294,316
Single-banked firms	6,619	2,608	21,493	23	171	431	877	1,942	7,773	1,510,737
Multiple-banked firms	26,244	9,505	70,824	42	319	847	1,836	4,722	27,051	4,294,316
Unbanked	3,722	1,599	3,942	11	125	341	718	1,557	5,622	124,397

[1]Spanish non financial firms: public and listed firms excluded.

Table 9. Distribution of firms across industries in the groups of unbanked, single-banked and multiple-banked firms (%).

	single banked	multiple banked	unbanked
Agriculture and fishing	2.46	1.42	2.67
Manufacturing ans quarrying	31.36	41.13	26.62
Electricity, gas and water supply	0.52	0.27	0.69
Construction and real estate developers	12.81	12.55	14.46
Wholesale and retail trade	35.56	31.94	37.63
Hotels and restaurants	4.14	1.93	5.55
Transport, storage and commuynications	4.30	4.67	2.55
IT, R&D and others	5.05	4.16	6.30
Other non financial industries	3.81	1.94	3.54

[1]Spanish non-financial firms: public and listed firms excluded.
Annual averages for 1997-2004.

Table 10. Descriptive statistics for the investment-to-capital ratio: all firms, unbanked firms, single-banked firms and multiple-banked firms (units per unit of capital).

	number of observations[1]	mean	standard deviation	min. value	p5	p25	p50	p75	p95	max. value
All firms	36,585	0.253	0.257	-0.575	0.000	0.062	0.185	0.387	0.776	1.407
Single-banked firms	6,619	0.237	0.262	-0.570	0.000	0.041	0.155	0.368	0.786	1.389
Multiple-banked firms	26,244	0.264	0.256	-0.575	0.000	0.075	0.200	0.400	0.778	1.404
Unbanked	3,722	0.204	0.250	-0.551	-0.006	0.020	0.121	0.321	0.728	1.388

[1]Spanish non financial firms: public and listed firms excluded.
Relative investment is defined as the ratio of (gross) investment on capital.
Annual averages for 1997-2004.

Table 11. Distribution of firms according to the sign of gross investment. All firms, unbanked firms, single-banked firms and multiple-banked firms (%).

	number of observations[1]	negative investment	zero investment	positive investment
All firms	36,585	4.39	3.89	91.73
Single-banked firms	6,619	4.42	6.42	89.17
Multiple-banked firms	26,244	4.20	2.16	93.64
Unbanked firms	3,722	5.65	11.50	82.85

[1]Spanish non-financial firms: public and listed firms excluded.
Annual averages for 1997-2004.

Table 12. Descriptive statistics for some economic and financial ratios (% and units per unit of capital).

	number of observations[1]	mean	standard deviation	min. value	p5	p25	p50	p75	p95	max. value
				single-banked firms						
Leverage ratio (%)	5,984	51.3	22.9	0.7	14.0	33.5	51.4	69.2	88.1	135.6
Bank debt-to-total debt ratio (%)	5,856	26.4	23.4	0.0	1.4	7.3	19.2	39.9	75.1	100.0
Liquidity ratio (%)	6,523	51.6	118.0	0.0	0.0	7.3	23.4	55.1	174.1	2280.0
Cash flow-to-capital ratio	6,612	0.8	1.2	-1.7	0.0	0.2	0.4	0.9	2.9	11.8
Sales-to Capital ratio	6,610	14.1	21.8	0.1	0.4	2.6	6.4	15.8	54.5	225.3
Return on Assets (%)	6,612	7.8	9.2	-53.2	-3.6	2.8	6.3	11.6	23.7	72.7
Default ratio* (%)	12	88.9	21.9	36.7	40.9	83.8	100.0	100.0	100.0	100.0
				multiple-banked firms						
Leverage ratio (%)	25,217	64.1	20.2	1.7	27.3	50.6	66.4	79.4	92.1	150.6
Bank debt-to-total debt ratio (%)	23,914	44.7	23.8	0.0	7.0	26.0	44.2	62.1	85.9	100.0
Liquidity ratio (%)	25,687	20.8	54.3	0.0	0.0	2.3	8.2	22.3	75.5	2280.0
Cash flow-to-capital ratio	26,230	0.7	1.0	-1.7	0.0	0.2	0.4	0.9	2.6	12.1
Sales-to Capital ratio	24,505	12.9	19.4	0.1	0.9	3.0	6.3	14.0	47.6	225.0
Return on Assets (%)	26,216	7.7	7.4	-55.2	-1.2	3.7	6.5	10.6	20.6	72.7
Default ratio* (%)	218	40.7	29.3	5.1	6.2	13.4	35.3	64.0	93.0	100.0
				unbanked firms						
Leverage ratio (%)	2,993	40.1	24.2	0.0	5.1	20.1	37.3	58.1	83.6	99.5
Bank debt-to-total debt ratio (%)	---	---	---	---	---	---	---	---	---	---
Liquidity ratio (%)	3,693	95.9	202.7	0.0	2.1	17.9	43.8	94.8	318.4	2280.0
Cash flow-to-capital ratio	3,713	0.9	1.4	-1.7	-0.1	0.2	0.5	1.1	3.5	11.9
Sales-to Capital ratio	3,710	17.1	24.7	0.1	0.3	3.0	8.5	20.5	63.9	218.9
Return on Assets (%)	3,717	8.7	10.6	-50.5	-3.8	2.8	6.9	13.2	27.3	72.7
Default ratio* (%)	---	---	---	---	---	---	---	---	---	---

[1]Spanish non-financial firms: public and listed firms excluded
* Subsample of firms in default.
Annual averages for 1997-2004.

Table 13. Arellano-Bond GMM parameter estimates of investment Euler equation [1]. All firms1997-2004.

Dependent variable: $(I/k)_t$

$(I/k)_{t-1}$	1.1414
	(0.000)
$(I/k)_{t-1}^2$	-0.8071
	(0.000)
$(q/k)_{t-1}$	0.0397
	(0.000)
$(y/k)_{t-1}$	0.0024
	(0.000)
AR(1)	-18.93
	(0.000)
AR(2)	10.94
	(0.000)
AR(3)	1.36
	(0.175)
Hansen-J test	68.68
	(0.008)
Number of observations	276,174
Number of firms	36,585

Parameter estimates correspond to equation:

$$\left(\frac{I}{k}\right)_{jt} = \beta_1\left(\frac{I}{k}\right)_{jt-1} + \beta_2\left(\frac{I}{k}\right)_{jt-1}^2 + \beta_3\left(\frac{q}{k}\right)_{jt-1} + \beta_4\left(\frac{y}{k}\right)_{jt-1} + d_t + \eta_j + \varepsilon_{jt},$$

where *I/k*, *q/k* and *y/k* denote respectively the ratios of gross investment, cash-flow and output to capital. d_t denotes time dummy variables.
List of instruments: third and fourth lags of *I/k, q/k* and *y/k.*
The number of observations corresponds to the selective panel dataset, over period 1997 to 2004.
P-values in parenthesis (two-sided).

Table 14. Arellano-Bond GMM parameter estimates of investment Euler equations across groups of firms with different numbers of banking relationships, 1997-2004.

Dependent variable: $(I/k)_t$

	(1)	(2)	(3)
Reference group (unbanked firms)			
$(I/k)_{t-1}$	1.9794	0.7149	1.0053
	(0.007)	(0.539)	(0.000)
$(I/k)_{t-1}^2$	-2.3307	-0.8832	-0.6544
	(0.003)	(0.492)	(0.016)
$(q/k)_{t-1}$	-0.1933	-0.1535	-0.1129
	(0.027)	(0.136)	(0.096)
$(y/k)_{t-1}$	0.0096	0.0080	0.0026
	(0.053)	(0.158)	(0.001)
Differences respect to the reference group			
Banked firms			
$(I/k)_{t-1}$	-1.5577		
	(0.069)		
$(I/k)_{t-1}^2$	2.3935		
	(0.012)		
$(q/k)_{t-1}$	0.2569		
	(0.010)		
$(y/k)_{t-1}$	-0.0090		
	(0.132)		
Single-banked firms			
$(I/k)_{t-1}$		-0.1901	
		(0.913)	
$(I/k)_{t-1}^2$		0.3350	
		(0.870)	
$(q/k)_{t-1}$		0.3038	0.3083
		(0.023)	(0.003)
$(y/k)_{t-1}$		-0.0043	
		(0.646)	
Multiple-banked firms			
$(I/k)_{t-1}$		0.1617	
		(0.908)	
$(I/k)_{t-1}^2$		0.5685	
		(0.707)	
$(q/k)_{t-1}$		0.1748	0.1157
		(0.185)	(0.141)
$(y/k)_{t-1}$		-0.0061	
		(0.367)	
AR(1)	-18.35	-8.18	-11.57
	(0.000)	(0.000)	(0.000)
AR(2)	6.7	3.5	4.88
	(0.000)	(0.000)	(0.000)
AR(3)	1.1	0.24	-0.23
	(0.273)	(0.812)	(0.822)
Hansen J test	39.85	21.95	29.18
	(0.432)	(0.740)	(0.658)
Number of observations	276,174	276,174	276,174
Number of firms	36,585	36,585	36,585

Parameter estimates in column (1) correspond to equation:

$$\left(\frac{I}{k}\right)_{jt} = \beta_1\left(\frac{I}{k}\right)_{jt-1} + \beta_2\left(\frac{I}{k}\right)^2_{jt-1} + \beta_3\left(\frac{q}{k}\right)_{jt-1} + \beta_4\left(\frac{y}{k}\right)_{jt-1} + d_t + \delta_1\left(\frac{I}{k}\right)_{jt-1} \times D0_{jt} + \delta_2\left(\frac{I}{k}\right)^2_{jt-1} \times D0_{jt} + \delta_3\left(\frac{q}{k}\right)_{jt-1} \times D0_{jt} + \delta_4\left(\frac{y}{k}\right)_{jt-1} \times D0_{jt} + d_t \times D0_{jt} + \eta_j + \varepsilon_{jt}.$$

where *I/k*, *q/k* and *y/k* denote respectively the ratios of gross investment, cash-flow and output to capital, d_t denotes time dummy variables and *D0* denotes a dummy variable that equals 1 if the firm is banked and 0 if the firm is unbanked (reference group). Analogously, parameter estimates in column (2) correspond to equation:

$$\left(\frac{I}{k}\right)_{jt} = \beta_1\left(\frac{I}{k}\right)_{jt-1} + \beta_2\left(\frac{I}{k}\right)^2_{jt-1} + \beta_3\left(\frac{q}{k}\right)_{jt-1} + \beta_4\left(\frac{y}{k}\right)_{jt-1} + d_t + \sum_{S=1,2}\left(\delta_1^S\left(\frac{I}{k}\right)_{jt-1} \times DS_{jt} + \delta_2^S\left(\frac{I}{k}\right)^2_{jt-1} \times DS_{jt} + \delta_3^S\left(\frac{q}{k}\right)_{jt-1} \times DS_{jt} + \delta_4^S\left(\frac{y}{k}\right)_{jt-1} \times DS_{jt} + d_t \times DS_{jt}\right) + \eta_j + \varepsilon_{jt},$$

where *D1* denotes a dummy variable that equals 1 if the firm is single-banked and 0 otherwise. *D2* denotes a dummy variable that equals 1 if the firm is multiple-banked and 0 otherwise. Consequently unbanked firm are the reference group. Parameter estimates in column (3) correspond to a restricted version of this equation that imposes $\delta_l^s = 0$, for $l=1, 2, 4$ and $s=1,2$.

List of instruments: third and fourth lags of *I/k*, *q/k* and *y/k* and third lag of the number of bank relations.

The number of observations corresponds to the selective panel dataset, over period 1997 to 2004.

P-values in parenthesis (two-sided).

Table 15. Arellano-Bond GMM parameter estimates of investment Euler equation: continuing unbanked firms, continuing single-banked firms continuing multiple-banked firms and rest of the firms

Dependent variable: $(I/k)_t$

	continuing unbanked	continuing single-banked	continuing multi-banked	rest of the firms
$(I/k)_{t-1}$	0.9976	-1.1011	0.1369	1.8415
	(0,539)	(0.733)	(0.705)	(0.000)
$(I/k)_{t-1}^2$	0.2283	0.3382	0.5225	-1.9669
	(0,907)	(0.917)	(0.271)	(0.002)
$(q/k)_{t-1}$	-0.3109	0.0715	-0.0169	0.1535
	(0,076)	(0.759)	(0.758)	(0.014)
$(y/k)_{t-1}$	0.0179	-0.0119	0.0012	0.0036
	(0,201)	(0.596)	(0.746)	(0.350)
AR(1)	-10.89			
	(0,000)			
AR(2)	4.74			
	(0,000)			
AR(3)	0.26			
	(0,792)			
Hansen-J test	45.09			
	(0,232)			
Number of observations	276,174			
Number of firms	36,585			

Parameter estimates correspond to equation:

$$\left(\frac{I}{k}\right)_{jt} = \beta_1\left(\frac{I}{k}\right)_{jt-1} + \beta_2\left(\frac{I}{k}\right)^2_{jt-1} + \beta_3\left(\frac{q}{k}\right)_{jt-1} + \beta_4\left(\frac{y}{k}\right)_{jt-1} + \sum_{S=1,2,3}\left(\delta_1^S\left(\frac{I}{k}\right)_{jt-1} \times DS_{jt} + \delta_2^S\left(\frac{I}{k}\right)^2_{jt-1} \times DS_{jt} + \delta_3^S\left(\frac{q}{k}\right)_{jt-1} \times DS_{jt} + \delta_4^S\left(\frac{y}{k}\right)_{jt-1} \times DS_{jt}\right) + d_t + \eta_j + \varepsilon_{jt},$$

where *I/k*, *q/k* and *y/k* denote respectively the ratios of gross investment, cash-flow and output to capital and d_t denotes time dummy variables. *D1* denotes a dummy variable that equals 1 if the firm is continuing single-banked and 0 otherwise. *D2* equals 1 if the firm is continuing multiple-banked and 0 otherwise. *D3* equals 1 if the firm belong to the category of rest of the firms and 0 otherwise.

Coefficients reported in the table are: for continuing unbanked firms, $\beta_l,\ \ l = 1,...,4$; for continuing single banked firms. $\beta_l + \delta_l^1,\ \ l = 1,...,4$; for continuing multi-banked firms, $\beta_l + \delta_l^2,\ \ l = 1,...,4$; and for rest of the firms, $\beta_l + \delta_l^3,\ \ l = 1,...,4$.

[a] List of instruments. 3rd and 4th lags of *I/k*, *q/k* and *y/k* and 3rd lag of the number of bank relations.

P-values in parenthesis (two-sided).

The number of observations corresponds to the selective panel dataset, over period 1997 to 2004.

Table 16. Baseline financial regime comparisons, Voung test Z-statistics[1,2,3]

Comparison	MH v FI	MH v B	MH v A	FI v B	FI v A	B v A	Best Fit
1. using 1997 (k,I,q) data							
-whole sample, n = 36,585	0.00(tie)	-62.4***(B)	46.5***(MH)	-62.4***(B)	46.5***(FI)	81.7***(B)	B
-continuously unbanked, n = 1,993	7.06***(MH)	4.86***(MH)	16.3***(MH)	1.63(tie)	15.4***(FI)	16.1***(B)	MH
-continuously single-banked, n = 914	0.24(tie)	-10.6***(B)	13.8***(MH)	-10.6***(B)	13.8***(FI)	13.1***(B)	B
-continuously multi-banked, n = 19,0	-45.5***(FI)	-33.4***(B)	4.31***(MH)	-21.3***(B)	34.8***(FI)	27.6***(B)	B
-others, n = 14,595	-38.9***(FI)	-50.9***(B)	10.1***(MH)	-44.1***(B)	41.8***(FI)	54.0***(B)	B
2. using 2004 (k,I,q) data							
-whole sample, n = 36,585	56.1***(MH)	-41.3***(B)	9.48***(MH)	-41.3***(B)	9.48***(FI)	51.6***(FI)	B
-continuously unbanked, n = 1,993	1.88*(MH)	14.5***(MH)	14.4***(MH)	14.5***(FI)	14.4***(FI)	0.00(tie)	MH
-continuously single-banked, n = 914	0.06(tie)	-6.56***(B)	13.3***(MH)	-6.56***(B)	13.3***(FI)	8.81***(B)	B
-continuously multi-banked, n = 19,0	-25.0***(FI)	-22.8***(B)	0.00(tie)	-8.89***(B)	25.0***(FI)	22.8***(B)	B
-others, n = 14,595	-6.99***(FI)	-30.7***(B)	51.8***(MH)	-28.5***(B)	46.3***(FI)	42.0***(B)	B

[1]***=1%, **=5%, *=10% two-sided significance level, the "winning" regime is in the parentheses.
[2] Vuong Z-statistics' cutoffs: 2.575 = *** 1.96 = ** 1.645 = * "tie"
[3] All results computed under the assumption of risk neutrality (σ=0).
B=borrowing or lending in non-contingent debt; MH=moral hazard constrained credit/insurance; FI=complete markets/full information; A=no access to credit (autarky).

Table 17. Parameter estimates using 1997 data, risk-neutral[1].

Model	γ_{me}	θ	ρ	$\mu_{w/b}$	$\gamma_{w/b}$	LL Value[2]
continuously unbanked firms						
Moral Hazard, MH *	0.2169 *(.004)*	20.833 *(.041)*	-1.5464 *(.000)*	0.7404 *(.000)*	0.0043 *(.000)*	**-4.7801**
Full Information, FI	0.2179 *(.004)*	53.106 *(30.3)*	-1.5469 *(.000)*	0.8301 *(.129)*	0.1006 *(.026)*	**-4.7885**
Borrowing or lending, B	0.2194 *(.004)*	12.806 *(.001)*	-1.5464 *(.184)*	0.2000 *(.000)*	0.0272 *(.004)*	**-4.7910**
Autarky, A	0.2870 *(.005)*	0.5000 *(.000)*	182.91 *(2.07)*	n.a.	n.a.	**-4.9230**
continuously single-banked firms						
Moral Hazard, MH	0.2396 *(.011)*	54.682 *(.000)*	-1.5463 *(.000)*	0.0960 *(.000)*	0.0000 *(.000)*	**-5.0563**
Full Information, FI	0.2396 *(.011)*	2.0283 *(.000)*	-1.5463 *(.000)*	0.2000 *(.000)*	0.0000 *(.000)*	**-5.0563**
Borrowing or lending, B *	0.0934 *(.006)*	0.003 *(.000)*	-0.0535 *(.053)*	0.8776 *(.010)*	0.1400 *(.010)*	**-4.7070**
Autarky, A	0.3245 *(.013)*	0.5000 *(.000)*	177.04 *(63.7)*	n.a.	n.a.	**-5.2364**
continuously multiple-banked firms						
Moral Hazard, MH	0.2732 *(.004)*	16.798 *(.000)*	0.3029 *(.000)*	0.9395 *(.000)*	0.0010 *(.000)*	**-5.4781**
Full Information, FI	0.2478 *(.003)*	17.040 *(.000)*	0.3029 *(.000)*	0.0616 *(.000)*	0.0019 *(.000)*	**-5.3941**
Borrowing or lending, B *	0.1332 *(.001)*	0.0027 *(.000)*	-0.4140 *(.000)*	0.9031 *(.002)*	0.0937 *(.003)*	**-5.2418**
Autarky, A	0.2944 *(.003)*	5.0000 *(.000)*	1088.3 *(.000)*	n.a.	n.a.	**-5.4935**
other firms						
Moral Hazard, MH	0.2859 *(.026)*	10.276 *(3.59)*	454.99 *(458)*	0.0295 *(.011)*	0.7240 *(.311)*	**-5.2157**
Full Information, FI	0.2282 *(.003)*	5.123 *(2.56)*	-1.0089 *(.150)*	0.0160 *(.049)*	0.0218 *(.066)*	**-5.0847**
Borrowing or lending, B *	0.1038 *(.001)*	0.0027 *(.000)*	-0.4140 *(.000)*	0.9152 *(.003)*	0.3199 *(.077)*	**-4.7528**
Autarky, A	0.3149 *(.004)*	0.5000 *(.000)*	177.04 *(.000)*	n.a.	n.a.	**-5.2372**

[1] Bootstrap standard errors in the parentheses.
[2] Normalized (divided by *n*) log-likelihood values
[3] $\mu_{w/b}$ and $\gamma_{w/b}$ reported relative to grid span.
*denotes the best fitting regime (including ties).

Table 18. Financial regime comparisons - Robustness, Voung test Z-statistics

Comparison	MH v FI	MH v B	MH v A	FI v B	FI v A	B v A	Best Fit
1. allowing for quadratic capital adjustment costs							
-continuously unbanked	0.56(tie)	4.43***(MH)	14.2***(MH)	-0.11(tie)	9.88***(FI)	13.2***(B)	MH,FI
-continuously single-banked	-3.45***(FI)	-10.0***(B)	15.0***(MH)	-9.17***(B)	13.2***(FI)	12.5***(B)	B
-continuously multi-banked	-10.7***(FI)	-23.3***(B)	26.9***(MH)	-10.6***(B)	32.7***(FI)	27.9***(B)	B
-others	-44.2***(FI)	-44.6***(B)	11.8***(MH)	-33.8***(B)	46.4***(FI)	46.3***(B)	B
2. allowing for risk aversion							
-continuously unbanked	12.4***(MH)	11.0***(MH)	16.0***(MH)	-1.75*(B)	10.6***(FI)	18.2***(B)	MH
-continuously single-banked	11.3***(MH)	-2.65***(B)	14.2***(MH)	-10.6***(B)	11.9***(FI)	12.5***(B)	B
-continuously multi-banked	39.1***(MH)	-26.7***(B)	64.0***(MH)	-60.6***(B)	34.8***(FI)	73.6***(B)	B
-others	42.7***(MH)	-12.8***(B)	47.7***(MH)	-44.9***(B)	24.3***(FI)	45.1***(B)	B
3. alternative grids							
-continuously unbanked	0.04(tie)	2.76***(MH)	2.48**(MH)	2.81***(FI)	2.52**(FI)	0.00(tie)	MH,FI
-continuously single-banked	2.72***(MH)	-4.39***(B)	2.84***(MH)	-3.33***(B)	0.00(tie)	3.10***(B)	B
-continuously multi-banked	0.00(tie)	-17.1***(B)	-91.2***(A)	-17.1***(B)	-83.7***(A)	17.1***(B)	B
-others	0.11(tie)	-8.10***(B)	7.95***(MH)	-8.10***(B)	7.95***(FI)	10.6***(B)	B

[1] ***=1%, **=5%, *=10% two-sided significance level, the "winning" regime is in the parentheses.
[2] Voung Z-statistics' cutoffs: 2.575 = *** 1.96 = ** 1.645 = * "tie"
[3] All results are computed using 1997 (k, I, q) data.

Table 19. Decomposition of aggregate debt[1] (%).

	Banked firms	Unbanked firms
Bank debt	33.9	0.0
Debt with associated and affiliated companies	17.1	71.8
Trade credit	28.6	10.5
Government	9.5	3.5
Commercial paper	3.4	0.3
Deposits and guarantees	0.6	0.4
Accounts receivable	2.1	1.0
Deffered debts on shares	0.1	0.2
Provisions for current assets	1.2	1.5
Bonds	0.1	0.0
Other debts*	3.3	10.8
Aggregate short and long term debt	100	100
Sample size	2,220	248

[1] Statistics based on the sample of firms that report to the Spanish Mercantile Registry detailed financial statements which contain the itemization of debt (2,468 firms). Specifically, figures in this table refer to 2,220 firms reporting positive banking debt ("banked") and 248 firms reporting no borrowing from banks ("unbanked"). All information refers to 1997.
* For banked firms, this category includes an adjustment of -2 percentage points which allows that the sum of items equals 100% of debt. For unbanked firms, it includes unexplained aggregate debt that represents 0.6% of total debt.

APPENDIX 1: Allowing for transitions across financial regimes

We show how we can extend our baseline model from Section 2 to allow transitions across financial regimes. The Euler equation approach is already consistent with this, in the sense that the Lagrange multiplier λ_t^D can vary with *t*, including taking on a value of zero. That is, the firm can be constrained at some times and not at other times. As noted in the two-period example in Section 2, there is a tendency for this to happen over time as small firms pay out no dividends and have λ_t^D positive, while large unconstrained firms have $\lambda_t^D = 0$. Likewise, the Karaivanov and Townsend (2010) approach can be also generalized to allow for transitions.

More generally, imagine that there is a fixed cost $f^B(\theta_t^B)$ for banking with a formal financial institution and cost $f^U(\theta_t^U)$ for being "unbanked". In some situations, firms have less reason to borrow from the formal sector as they get larger, because their needs are smaller, as illustrated earlier. In this case, with $f^U = 0$ and $f^B > 0$, they exit the banked sector at some point. In the extreme, the costs $f^U(\theta_t^U)$ may be persistently low relative to $f^B(\theta_t^B)$ and the firm will never be observed to be borrowing from the formal sector.

On the other hand, there could be a cost of exiting the banking system or, equivalently, cost to joining an alternative financial arrangement. More generally, shocks may cause firms to transition back and forth over time. Nevertheless, one can formalize the firm's problem. Returning to our earlier borrowing model, if a firm remains in the banking system at *t*, and faces that decision again next period, it solves

$$W^B\left(k_{t-1}, \theta_t^q, B_{t-1}, S_{t-1,} \theta_t^I\right) = \max_{I_t, B_t, S_t} c_t + \beta E \max\left[W^U(.), W^B(.)\right]$$

where

$$c_t = k_{t-1}(1-\delta) + q(k_t, \bar{z}_t, \theta_t^q) - k_t - R^B B_{t-1} + B_t + R^S S_{t-1} - S_t - \Psi_B(B_t) - g(k_t, I_t, \theta_t^I) - f^B(\theta_t^B)$$

If when transitioning at time *t* to regime "*U*" (unbanked) previous debt obligations are met, then only current investment need be decided:

$$W^U\left(k_{t-1}, \theta_t^q, B_{t-1}, S_{t-1,} \theta_t^I\right) = \max_{I_t} c_t + \beta E \max\left[W^U(.), W^B(.)\right],$$

where

$$c_t = k_{t-1}(1-\delta) + q(k_t, \bar{z}_t, \theta_t^q) - k_t - R^B B_{t-1} + R^S S_{t-1} - g(k_t, I_t, \theta_t^I) - f^U\left(\theta_t^U\right),$$

and where $k_t = k_{t-1}(1-\delta) + I_t$ throughout.

Though there may be long-run tendencies to move toward the unbanked state, shocks to profits, adjustment costs and entry or exit costs will cause some transitions. On the other hand, if those transitions are distant in expectation and in time, the value functions W^B and W^U can be approximated by imagining that the firm stays in one status or another forever (see Townsend and Ueda, 2007). This is what we have in mind empirically when using the continuing stratifications in the MLE approach.

APPENDIX 2: Variables definitions

Firm characteristics

-*Age* (relative to the constitution date).

-*Industry* (3 and 4 digit CNAE-93 classifications). Financial firms have been excluded. Using this information, we have aggregated the activities into: 1) agriculture and fishing; 2) manufacturing and quarrying; 3) electricity, gas and water supply; 4) construction and property development; 5) wholesale and retail trade; 6) hotels and restaurants; 7) transport, storage and communication; 8) IT, R&D and other; 9) other non-financial industries. We build up a set of dummy variables controlling for heterogeneity across industries in empirical models. Using the 3 digit CNAE-classification would imply estimating a large number of parameters which is not viable because of capacity constraints.

-*Location* (province).

Firm-level economic variables

-*Capital* (k_t): The stock of capital in year *t* is measured by the book value of tangible fixed assets. This variable has been deflated using the gross fixed capital investment deflator.

-*(Gross) Investment* (I_t): Defined as the sum of the absolute increase in the stock of tangible fixed assets between *t* and *t-1* and the depreciation in year *t*. Deflated using the gross fixed capital investment deflator.

-*Cash flow* (q_t)*:* It is defined as the sum of operating profits/losses (+/-) and depreciation. Deflated using the GDP deflator.

-*Output* (y_t): It is measured using total sales. Deflated using the GDP deflator.

Economic and financial ratios

-*Relative Investment* (*I/k*): defined as the ratio of gross investment (*I*) to capital (*k*). The univariate series of relative investment has been truncated dropping out the 5% of observations contained in the upper and the lower tails of the distribution, which will greatly distort the results of the estimation of Euler equations. This data treatment is applied before picking up firms in the selective sample.

-*ROA*: Profitability ratio, profits before interests and taxes divided by total assets (winsorized 1%, before picking up firms in the selective sample).

-*Leverage* (ratio): Amount of debt with explicit cost divided by total assets (winsorised 1%, before picking up firms in the selective sample).

-*Liquidity* (ratio): cash to short term debt ratio (winsorised 1%, before picking up firms in the selective sample).

Firm credit risk

-*Banking debt*: amount of debt contracted with any Spanish credit institution. Aggregation, at firm level, of the amounts of banking debt registered by Spanish credit Institutions in CIR.

-*Default banking debt***:** amount of non-performing loans (more than three month in arrears) and doubtful loans (low repayment probability despite of being performing). It is constructed by aggregating, at the firm level, the amounts of non-performing and doubtful loans registered by Spanish credit Institutions in CIR.

-*Firm-level default ratio*: defined as the proportion of banking debt that is non-performing or doubtful, if this proportion exceeds 5%. If it is below this threshold, the default ratio is assumed equal to zero and default banking debt is imputed as zero.[64]

-*Default dummy*: dichotomic variable that equals one if firm-level default ratio is higher than 5% and zero otherwise.

Firm's banking relationships

-*Number of relations*: number of different Spanish credit institutions granting a loan to the firm. According to the current number of relations, we distinguish year by year three types of firms: *unbanked*, in *single banking relationship* and in *multiple banking relationships*. According to the trajectories across these three categories over the whole period 1997-2004, we distinguish among *continuing unbanked*, *continuing banked* and *switching firms*. Within the category of continuing banked firms we differentiate among *continuing single-banked*, *continuing multiple-banked* and *other continuing banked firms*. Empirically, for analysing differences across groups of firms defined according to their trajectories in and out of the unbanked status, we consider for categories: *continuing unbanked*, *continuing single-banked*, *continuing multiple-banked* and *rest of the firms* (included both switching firms and other continuing banked firms).

[64] The minimum threshold of 5% allows filtering "technical defaults" and mistakes, which are unrelated to the level of credit risk of the firms. This filter has been applied in other papers that use the CIR database (e.g. Jiménez and Saurina (2004)).

APPENDIX 3: Descriptive statistics based on the unbalanced panel dataset

Table A3.1 exhibits the sample sizes for every cross-section in the SABI-CIR database. Additionally, it shows some statistics about the coverage of the database relative to the population of Spanish non-financial firms in terms of the aggregate figures of total assets and banking loans. On average, the gross number of firms in the sample is over 297,000 firms per year. The coverage of SABI-CIR has tended to increase over time, from a minimum of around 32%, in terms of assets, and 36.4%, in terms of banking debt, in 1994, to a percentage that reaches a maximum close to 72% of total assets and 71.3% of aggregate banking debt in Spanish non-financial firms, in year 2002. This increase in the coverage is explained by the expansionary strategy of INFORMA that has widely increased the number of firms for which it buys public statements deposited at the Mercantile Registry.

SABI-CIR database has a wide heterogeneity of the sample in terms of firms' age, activity and size. In terms of age, the sample is mostly composed of firms not older than 10 years, which exceed 63%. In this category, firms are concentrated to a large extent below 5 years old (34.2%). Retail, construction, real estate development and manufacturing are the industries in which three-quarters of all firms are concentrated. In terms of firms' size, measured by total assets, the database contains a high proportion of small firms. In particular, an average percentage of 93.4% of the firms in the sample correspond to the smallest firms' category, which includes firms with an amount of total asset not higher than 5 million € or with total sales not higher than 7 million €. This characteristic differentiates this dataset from alternative datasets relative to Spanish non-financial firms, which tend to under-represent small firms (see Tables A3.2, A3.3 and A3.4).

A comparison of the selective and the unbalanced samples

Table A3.5 shows the distribution of firms across the categories of small-size, medium-size and large-size over the period 1997-2004.[65] As in the unbalanced sample, most of the firms in the selective sample belong to the category "small firms", with an average weight of 83%. Still, compared with the composition of the unbalanced sample, the selective sample tends to over-represent the categories of medium-sized firms and large-sized firms, with average fractions of 14.9% and 2.1%, respectively, which are more than twice the corresponding percentages for the entire SABI-CIR sample.

Additionally, as shown in Table A3.6, our sample is also more homogeneous than the population in terms of firm age. On average, more than three quarters of the firms in the sample[66] are older than ten years. Moreover, the homogeneity of the sample in terms of age is increasing over time, with the percentage of firms in the category of firms younger than 10 years being below 4%. In the entire SABI-CIR sample, firms are almost equally distributed across the three age categories (Table A3.2) over the analysed period. This is explained by

[65] Firm size classification follows EU recommendations (Bulletin EU 1/2-1996), which defines as small firms those with total assets lower than 5 million Euros or sales lower than 7 million Euros and as large firms those with total assets higher than 27 million Euros or sales higher than 40 million Euros.

[66] The statistics about the age distribution of the firms in the sample are computed from the subsample of firms with data available on the age of the firm, which reaches only 40% of the firms in the sample.

the exclusion of new cohorts of entering firms as a consequence of the selection criteria requiring complete information for period 1997-2004.

The higher homogeneity of the sample of firms included in the selective panel is also evident in the distribution of firms across industries.[67] Statistics presented in A3.7 indicate that firms in the selective sample are concentrated in the industries of *Manufacturing and Quarrying*; *Wholesale and Retail Trade* and *Construction and Real Estate Developers* to a larger extent than in the non-selective sample of firms. These three industries concentrate 84% of the firms in the sample against the 74% in the non-selective sample. The rest of the firms are spread over the rest of industries.

Table A3.1. Coverage of the SABI-CIR sample of private and unlisted firms (million € and %).

	sample size[1]	coverage in terms of:	
		bank loans[2]	total assets[3]
1993	101,290	44.5	
1994	96,519	36.4	32.3
1995	170,346	58.9	51.6
1996	214,449	64.2	54.1
1997	224,811	66.3	61.9
1998	266,059	67.1	56.3
1999	306,453	66.3	60.2
2000	341,350	67.0	64.1
2001	429,952	69.8	67.3
2002	505,235	71.3	71.9
2003	515,722	72.0	68.8
2004	400,299	60.9	58.4

1. Private and non-listed Spanish non-financial firms (after filtering).
2. Coverages in terms of bank loans are relative to the aggregated values of banking debt for Spanish non-financial firms (public and listed firms included) computed using CIR data.
3. Coverages have been computed over the aggregated value of total assets corresponding to Spanish non-financial firms (public and listed firms included) published in the annual reports of Central de Balances (Banco de España).

Table A3.2. Distribution of firms in the SABI-CIR database across age categories

	1997	1998	1999	2000	2001	2002	2003	2004	average 1997-2004
[0, 5)	36.8	37.4	35.6	33.2	33.9	34.5	33.3	29.3	34.2
[5-10)	27.9	26.9	28.2	30.5	31.3	30.5	30.2	28.7	29.3
[10,∞]	35.4	35.7	36.1	36.4	34.8	35.0	36.5	42.0	36.5
Number of non financial firms[1]	91,507	109,114	126,401	141,344	179,170	210,905	215,240	167,515	155,150

[1]Spanish non-financial firms: public and listed firms excluded.
[2]Statistics presented in this table are computed after excluding firms with no data available on age.

Table A3.3. Distribution of firms in the SABI-CIR database across industries

	1997	1998	1999	2000	2001	2002	2003	2004	Average 1997-2004
Agriculture and fishing	2.39	2.30	2.23	2.29	2.54	2.57	2.50	2.16	2.37
Manufacturing ans quarrying	23.77	22.02	20.87	20.08	18.35	16.71	15.99	16.84	19.33
Electricity, gas and water supply	0.37	0.37	0.37	0.38	0.37	0.36	0.37	0.37	0.37
Construction and real estate developers	23.81	24.28	25.21	25.94	26.87	28.42	29.31	28.83	26.59
Wholesale and retail trade	30.44	29.72	29.25	28.60	27.31	25.79	25.08	26.13	27.79
Hotels and restaurants	2.64	3.09	3.55	3.83	4.40	4.59	4.37	4.12	3.82
Transport, storage and commuynications	4.95	4.99	4.92	4.86	4.69	4.52	4.39	4.34	4.71
IT, R&D and others	8.48	9.68	9.91	10.10	11.08	12.20	13.05	12.81	10.91
Other non financial industries	3.15	3.54	3.68	3.91	4.40	4.84	4.94	4.41	4.11
Number of non financial firms[1]	224,120	265,647	306,074	341,082	429,439	504,586	515,221	399,960	373,266

[1]Spanish non-financial firms: public and listed firms excluded.

[67] Classification of firms across industries according to the information provided by INFORMA is static in the sense that it provides a unique CNAE code for every firm that does not vary over time.

Table A3.4. Distribution of firms in the SABI-CIR database across size categories[2]

	1997	1998	1999	2000	2001	2002	2003	2004	average 1997-2004
small-sized firms	93.5	93.6	93.5	93.2	93.8	94.1	93.6	92.3	93.4
medium-sized firms	5.9	5.8	5.8	6.1	5.6	5.3	5.8	6.9	5.9
large-sized firms	0.7	0.6	0.6	0.7	0.6	0.6	0.6	0.8	0.7
Number of non financial firms[1]	224,811	266,059	306,453	341,350	429,952	505,235	515,716	400,299	373,734

[1]Spanish non-financial firms: public and listed firms excluded.
[2]Small-sized firms: sales not higher than €7 million or total assets not higher than €5 million; medium-sized firms: sales higher than €7 million but not higher than €40 million or total assets higher than €5 million but not higher than €27 million; large-sized firms: sales higher than €40 million or total assets higher than €27 million.

Table A3.5. Distribution of firms in the selective panel across size categories[2] (%)

	1997	1998	1999	2000	2001	2002	2003	2004	average 1997-2004
small-sized firms	88.2	86.8	85.1	83.4	82.1	80.9	79.5	78.0	83.0
medium-sized firms	10.4	11.6	13.1	14.5	15.6	16.7	17.9	19.1	14.9
large-sized firms	1.4	1.6	1.8	2.1	2.3	2.4	2.6	2.9	2.1
Number of non financial firms[1]	44,644	44,644	44,644	44,644	44,644	44,644	44,644	44,644	44,644

[1]Spanish non-financial firms: public and listed firms excluded.
[2]Small-sized firms: sales not higher than €7 million or total assets not higher than €5 million; Medium-sized firms: sales higher than €7 million but not higher than €40 million or total assets higher than €5 million but not higher than €27 million; Large-sized firms: sales higher than €40 million or total assets higher than €27 million.

Table A3.6. Distribution of firms across age categories in the selective panel (%)

	1997	1998	1999	2000	2001	2002	2003	2004	average 1997-2004
[0, 5)	15.2	9.8	3.6	0.1	0.0	0.0	0.0	0.0	3.6
[5-10)	29.5	28.7	28.7	27.1	21.1	15.2	9.8	3.6	20.5
[10,∞]	55.3	61.5	67.7	72.9	78.8	84.8	90.2	96.4	75.9
Number of non financial firms[1]	18,009	18,015	18,015	18,015	18,015	18,015	18,015	18,015	18,014

[1]Spanish non-financial firms: public and listed firms excluded. Subsample of firms with age data.

Table A3.7. Distributions of firms across industries in the selective panel (%)

Agriculture and fishing	1.71
Manufacturing ans quarrying	36.75
Electricity, gas and water supply	0.33
Construction and real estate developers	13.26
Wholesale and retail trade	34.00
Hotels and restaurants	2.68
Transport, storage and commuynications	4.29
IT, R&D and others	4.57
Other non financial industries	2.42
Number of non financial firms[1]	44,644

[1]Spanish non-financial firms: public and listed firms excluded.

APPENDIX 4: Robustness checks: Blundell and Bond GMM estimates

This appendix provides in tables A4.1 and A4.2 robustness checks of the estimations reported in tables 14 and 15, respectively, using the Blundell and Bond (1998) extended first-differenced GMM estimator for investment Euler equations.

Table A4.1. Blundell-Bond GMM parameter estimates of investment Euler equation. All firms. Period 1997-2004.

Dependent variable: $(I/k)_t$

$(I/k)_{t-1}$	1.1114
	(0,000)
$(I/k)_{t-1}^2$	-0.6035
	(0,000)
$(q/k)_{t-1}$	0.0224
	(0,000)
$(y/k)_{t-1}$	0.0013
	(0,000)
AR(1)	-44.39
	(0,000)
AR(2)	23.58
	(0,000)
AR(3)	1.51
	(0,132)
Hansen-J test	137.15
	(0,000)
Number of observations	292,680
Number of firms	36,585

Parameter estimates correspond to equation:

$$\left(\frac{I}{k}\right)_{jt} = \beta_1\left(\frac{I}{k}\right)_{jt-1} + \beta_2\left(\frac{I}{k}\right)^2_{jt-1} + \beta_3\left(\frac{q}{k}\right)_{jt-1} + \beta_4\left(\frac{y}{k}\right)_{jt-1} + d_t + \eta_j + \varepsilon_{jt},$$

where *I/k*, *q/k* and *y/k* denote respectively the ratios of gross investment, cash-flow and output to capital. d_t denotes time dummy variables.

List of instruments:

- For first differenced equation: third and fourth lags of *I/k, q/k* and *y/k*.
- For equations in levels: third lag of first-differenced *I/k, q/k, y/k*.

The number of observations corresponds to the selective panel dataset, over period 1997 to 2004.

P-values in parenthesis (two-sided).

Table A4.2. Blundell-Bond GMM parameter estimates of Investment Euler equations across groups of firms with different numbers of banking relationships. Period 1997-2004.

Dependent variable: $(I/k)_t$	(1)	(2)	(3)
Reference group (unbanked firms)			
$(I/k)_{t-1}$	0.7133	1.3632	1.1183
	(0.072)	(0.023)	(0.000)
$(I/k)_{t-1}^2$	-0.5789	-1.1068	-0.6175
	(0.270)	(0.123)	(0.000)
$(q/k)_{t-1}$	-0.0631	-0.0340	-0.0835
	(0.121)	(0.454)	(0.004)
$(y/k)_{t-1}$	0.0016	0.0014	0.0008
	(0.477)	(0.591)	(0.000)
Differences respect to the reference group			
Banked firms			
$(I/k)_{t-1}$	0.4726		
	(0.270)		
$(I/k)_{t-1}^2$	-0.0746		
	(0.895)		
$(q/k)_{t-1}$	0.1059		
	(0.033)		
$(y/k)_{t-1}$	-0.0006		
	(0.840)		
Single-banked firms			
$(I/k)_{t-1}$		-1.0933	
		(0.238)	
$(I/k)_{t-1}^2$		1.3117	
		(0.244)	
$(q/k)_{t-1}$		0.0411	0.0790
		(0.520)	(0.109)
$(y/k)_{t-1}$		0.0028	
		(0.505)	
Multiple-banked firms			
$(I/k)_{t-1}$		-0.0351	
		(0.955)	
$(I/k)_{t-1}^2$		0.3345	
		(0.653)	
$(q/k)_{t-1}$		0.0791	0.1441
		(0.187)	(0.000)
$(y/k)_{t-1}$		-0.0015	
		(0.645)	
AR(1)	-47.61	-28.79	-39.05
	(0.000)	(0.000)	(0.000)
AR(2)	23.32	14.3	16.79
	(0.000)	(0.000)	(0.000)
AR(3)	1.96	0.042	0.42
	(0.050)	(0.672)	(0.677)
Hansen J test	192.65	131.2	174.96
	(0.000)	(0.001)	(0.000)
Number of observations	292,680	292,680	292,680
Number of firms	36,585	36,585	36,585

The parameter estimates in column (1) correspond to the equation:

$$\left(\frac{I}{k}\right)_{jt} = \beta_1\left(\frac{I}{k}\right)_{jt-1} + \beta_2\left(\frac{I}{k}\right)^2_{jt-1} + \beta_3\left(\frac{q}{k}\right)_{jt-1} + \beta_4\left(\frac{y}{k}\right)_{jt-1} + d_t + \delta_1\left(\frac{I}{k}\right)_{jt-1} \times D0_{jt} + \delta_2\left(\frac{I}{k}\right)^2_{jt-1} \times D0_{jt} + \delta_3\left(\frac{q}{k}\right)_{jt-1} \times D0_{jt} + \delta_4\left(\frac{y}{k}\right)_{jt-1} \times D0_{jt} + d_t \times D0_{jt} + \eta_j + \varepsilon_{jt}.$$

where *I/k*, *q/k* and *y/k* denote respectively the ratios of gross investment, cash-flow and output to capital; d_t denotes time dummy variables and *D0* denotes a dummy variable that equals 1 if the firm is banked and 0 if the firm is unbanked (reference group). Analogously, the parameter estimates in column (2) correspond to equation:

$$\left(\frac{I}{k}\right)_{jt} = \beta_1\left(\frac{I}{k}\right)_{jt-1} + \beta_2\left(\frac{I}{k}\right)^2_{jt-1} + \beta_3\left(\frac{q}{k}\right)_{jt-1} + \beta_4\left(\frac{y}{k}\right)_{jt-1} + d_t + \sum_{S=1,2}\left(\delta_1^S\left(\frac{I}{k}\right)_{jt-1} \times DS_{jt} + \delta_2^S\left(\frac{I}{k}\right)^2_{jt-1} \times DS_{jt} + \delta_3^S\left(\frac{q}{k}\right)_{jt-1} \times DS_{jt} + \delta_4^S\left(\frac{y}{k}\right)_{jt-1} \times DS_{jt} + d_t \times DS_{jt}\right) + \eta_j + \varepsilon_{jt},$$

where *D1* denotes a dummy variable that equals 1 if the firm is single-banked and 0 otherwise. *D2* denotes a dummy variable that equals 1 if the firm is multiple-banked and 0 otherwise. Consequently unbanked firm are the reference group. Parameter estimates in column (3) correspond to a restricted version of this equation that imposes β_l^s, for l=1, 2, 4 and s=1,2.

List of instruments:
- For first differenced equation: third and fourth lags of *I/k*, *q/k* and *y/k* and third lag of the number of bank relations.
- For equations in levels: third and fourth lags of first-differenced *I/k*, *q/k*, *y/k* and the number of bank relations.

The number of observations corresponds to the selective panel dataset, over period 1997 to 2004.
P-values in parenthesis (two-sided).

APPENDIX 5: Differences across firms with different trajectories in and out of the unbanked status: Euler equation approach

The Euler equations approach considers the potential endogeneity of the category to which the firm belongs, in addition to that of the rest of the variables on the right-hand side of the investment equation. Table A5 reports three robustness GMM estimations of the investment Euler equation alternative to those presented in Table 15 (section 4.1.2). Column (1) of Table A5 reports the Arellano and Bond GMM estimates of the investment Euler equation corresponding to these four categories. It relies on the orthogonality conditions defined by the 3rd and 4th lags of the endogenous and predetermined variables (*I/k*, *q/k* and *y/k*) without including lags in the banking status. Year by year, it may be reasonable to assume that the number of banking relations is simultaneously determined with the optimal investment rate. Consequently, lagged investment rates may in fact be instrumenting not only first-differences of investment rates but also past firm's decisions on the number of relations and, consequently, the continuing categories to which firms have been assigned to. The same applies to the lags of *q/k* and *y/k*.

Column (2) uses Blundell and Bond's approach to extend the orthogonality conditions of the Arellano and Bond estimator from column (1) to include the orthogonality between the error of the equation in levels and third lags of first-differenced *I/k*, *q/k*, *y/k*. Finally, column (3), in addition to these instruments, includes the orthogonality conditions between, first, the first differenced error and the third lag of the number of bank relations and, second, between the error of the equation in levels and the third lag of first differenced number of bank relations.

Table A5. GMM parameter estimates of investment Euler equation: continuing unbanked firms, continuing single-banked firms, continuing multiple-banked firms and rest of the firms.

Dependent variable: $(I/k)_t$

	(1)[a]	(2)[b]	(3)[c]
continuing unbanked			
$(I/k)_{t-1}$	0.2235	-0.0143	1.3798
	(0.928)	(0.992)	(0.153)
$(I/k)_{t-1}^2$	0.8637	0.0958	-1.6103
	(0.745)	(0.956)	(0.183)
$(q/k)_{t-1}$	-0.3484	-0.2501	-0.1088
	(0.110)	(0.069)	(0.192)
$(y/k)_{t-1}$	0.0161	0.0028	0.0023
	(0.311)	(0.715)	(0.669)
continuing single-banked			
$(I/k)_{t-1}$	1.8998	6.2955	4.9154
	(0.627)	(0.010)	(0.009)
$(I/k)_{t-1}^2$	-1.2114	-6.1195	-5.6755
	(0.738)	(0.020)	(0.006)
$(q/k)_{t-1}$	0.2754	0.1904	0.0145
	(0.358)	(0.348)	(0.923)
$(y/k)_{t-1}$	-0.0201	-0.0100	0.0067
	(0.442)	(0.553)	(0.565)
continuing multi-banked			
$(I/k)_{t-1}$	0.8933	1.0459	1.1049
	(0.212)	(0.003)	(0.000)
$(I/k)_{t-1}^2$	-0.5481	-0.6098	-0.5051
	(0.453)	(0.141)	(0.051)
$(q/k)_{t-1}$	-0.0592	0.1123	0.1069
	(0.953)	(0.014)	(0.001)
$(y/k)_{t-1}$	0.0003	-0.0051	-0.0036
	(0.009)	(0.118)	(0.073)
rest of the firms			
$(I/k)_{t-1}$	2.0936	0.9895	0.9701
	(0.009)	(0.032)	(0.000)
$(I/k)_{t-1}^2$	-1.9100	-0.5015	-0.6265
	(0.021)	(0.349)	(0.047)
$(q/k)_{t-1}$	0.1972	0.0170	0.0059
	(0.007)	(0.705)	(0.8543)
$(y/k)_{t-1}$	0.0051	0.0070	0.0053
	(0.308)	(0.006)	(0.003)
AR(1)	-8.21	-20.5	-29.71
	(0.000)	(0.000)	(0.000)
AR(2)	4.85	11.63	14.60
	(0.000)	(0.000)	(0.000)
AR(3)	0.02	0.98	1.76
	(0.982)	(0.326)	(0.078)
Hansen J test	23.73	55.57	137.26
	(0.821)	(0.415)	(0.000)
Number of observations	276,174	292,680	292,680
Number of firms	36,585	36,585	36,585

The above parameter estimates correspond to the equation:

$$\left(\frac{I}{k}\right)_{jt} = \beta_1\left(\frac{I}{k}\right)_{jt-1} + \beta_2\left(\frac{I}{k}\right)_{jt-1}^2 + \beta_3\left(\frac{q}{k}\right)_{jt-1} + \beta_4\left(\frac{y}{k}\right)_{jt-1} + \sum_{S=1,2,3}\left(\delta_1^S\left(\frac{I}{k}\right)_{jt-1}\times DS_{jt} + \delta_2^S\left(\frac{I}{k}\right)_{jt-1}^2\times DS_{jt} + \delta_3^S\left(\frac{q}{k}\right)_{jt-1}\times DS_{jt} + \delta_4^S\left(\frac{y}{k}\right)_{jt-1}\times DS_{jt}\right) + d_t + \eta_j + \varepsilon_{jt},$$

where *I/k*, *q/k* and *y/k* denote respectively the ratios of gross investment, cash-flow and output to capital and d_t denotes time dummy variables. *D1* denotes a dummy variable that equals 1 if the firm is continuing single-banked and 0 otherwise. *D2* equals 1 if the firm is continuing multiple-banked and 0 otherwise. *D3* equals 1 if the firm belong to the category of rest of the firms and 0 otherwise.

Coefficients reported in the table are: for continuing unbanked firms, $\beta_l,\ l=1,...,4$; for continuing single banked firms. $\beta_l+\delta_l^1,\ l=1,...,4$; for continuing multi-banked firms, $\beta_l+\delta_l^2,\ l=1,...,4$; and for rest of the firms, $\beta_l+\delta_l^3,\ l=1,...,4$.

[a] Arellano and Bond GMM estimator. List of instruments: 3rd and 4th lags of *I/k, q/k* and *y/k*.

[b] Blundell and Bond GMM estimator. List of instruments: for first differenced equation: 3rd and 4th lags of *I/k*, *q/k* and *y/k*; for equations in levels: 3rd lag of first-differenced *I/k*, *q/k* and *y/k*.

[c] Blundell and Bond GMM estimator. List of instruments: for first differenced equation: 3rd and 4th lags of *I/k*, *q/k* and *y/k* and 3rd lag of the *number of bank relations*; for equations in levels: 3rd lag of first-differenced *I/k*, *q/k*, *y/k* and the number of bank relations.

P-values in parenthesis (two-sided).

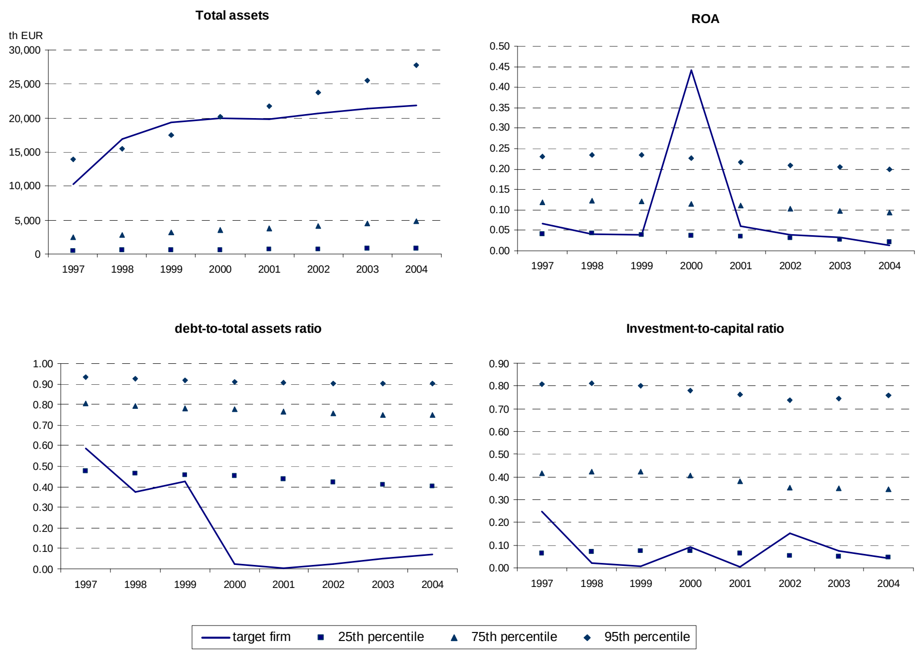
The number of observations corresponds to the selective panel dataset, over period 1997 to 2004.

APPENDIX 6: Case study of a large unbanked firm

Firm A is an example of a continuing unbanked firm that, as can be seen in Chart A6, is part of a relatively complex group of firms which are directly or indirectly owned by one person (*Mr. X*) and other members of the same family (wife, sons, daughters and mother).[68] To the best of our knowledge,[69] the structure is composed of 11 family firms, which are heterogeneous in terms of their age,[70] size[71] and activities.[72]

Figure A.6. provides in four separated panels the evolution over the period 1997-2004 of four key variables: total assets, Return on Assets (ROA), the investment-to-capital ratio, and the debt-to-total assets ratio for the target firm (*Firm A*) and also the 25th, 75th and 95th percentiles describing the distribution of these variables in the whole sample of firms (i.e., mature firms in the balanced or selective sample).

Figure A.6. Compared evolution of total assets, ROA, investment-to-capital ratio and debt-to-capital ratio in the target firm (Firm A) and in the 25th, 75th and 95th percentiles of the firm-level sample distributions of these variables



Firm A is a large firm in terms of total assets that exhibits an increasing trend over the analysed period. Compared to the rest of the firms in the sample under study, the firm is located in the upper tail of the

[68] In Spain, family names are composed by a first surname from the father and a second surname from the mother. This allows drawing conclusions about family relationships just from the direct comparison of the two surnames.

[69] We have built the structure of family related firms using SABI-INFORMA. This database provides for every firm a summary report that, among other information, contains the list of shareholders and the list of subsidiaries. When shareholders or subsidiaries are firms contained in the database, reports are linked. Thus, the whole structure can be built taking into account the information reported by "linked" firms. Nonetheless, this apparent structure would be incomplete because a given shareholder could own other "non-linked" firms, that is, firms that are not shareholders or subsidiaries. In this case, trying to discover relationships across firms is a more handcrafted task based on the study of clues such as similarities in firms' names, common addresses, fax or telephone numbers, etc. In this case, we have not identified more related firms.

[70] Constitution dates ranges from 1979 to 1994.

[71] Comparing the averages of the total assets in period 1997-2004 of every firm with the average percentiles of the distribution in the sample of total assets, two of them are below the 5th percentile, four between the 5th and the 25th percentile, two between the 25th and the 50th percentile and two between the 75th and the 95th percentile.

[72] According to our classification of industries: four of them are devoted to manufacturing (CNAE 2009 codes: 2599, 1103, 2599 and 2370); three to IT, R&D and others (CNAE 2009 codes: 7022), two to wholesale and/or retail trade (CNAE 2009 codes: 1979 and 4669); and two to agriculture (CNAE 2009 code 0161) and to financial intermediation (CNAE 2009 code 6492), respectively.

distribution. In particular, the level of assets of this firm is comparable to that of the 95th percentile of the firm level distribution of this variable over the entire analysed period.

The profitability of the studied firm is, with the exception of year 2000,[73] very close to the 25th percentile of the distribution of this variable in the sample. Leverage, measured by the ratio of debt to total assets, is low, particularly after a sharp decrease in 2000, from an average ratio of around 51% in 1997-1999, which is close to the 25th percentile of the sample distribution of firms' leverage ratio, to an average ratio of 3.5% in period 2000-2004, which is even lower than the 5th percentile (in average, it is around 18%). Finally, the average investment-to-capital ratio is 11.5% for the analysed firm which is about the 25th percentile of the sample distribution of this variable.

Firms related to Firm A are quite heterogeneous in terms of the weight of shareholders' funds as a percentage of total assets and also in terms of the composition of their liabilities. As shown in Table A.6.1 the average ratio of shareholders' funds-to-total assets varies across firms, from 26.1% to 100%. Nonetheless, except in one case (where it is 26.1%), the shareholders' funds ratio is higher than the (average) 75th percentile of the firm-level sample distribution of this ratio. Consequently, it can be concluded that the firms in the structure are highly capitalised. In aggregated terms (that is, for the pseudo-firm that results from aggregating all the firms in the structure), the average shareholders' fund ratio is above 70%.

These high solvency ratios are associated with a low dependence on banking debt. The numbers reported in Table A.6.2 indicate that taking all the firms in the structure as a group, banking debt represents, on average, approximately 2% of total assets. This aggregated figure is the result of a combination of firms that rely on banks' loans in different degrees. Most of them (seven) are always (or in almost all years) unbanked during the period. Some of them are continuously (two) or intermittently (two) banked. It is remarkable that, within the subset of banked firms, bank debt as a percentage of total assets does greatly vary across firms and even for the same firm over the years. Firms with higher banking debt ratios compared with the rest of the firms in the structure (i.e., the cases of *Firm K* and, in year 2004, *Firm J*) also exhibit the relatively lower shareholders' ratios.

Table A.6.1. Shareholders' funds-to-total assets ratio (%)

	Shareholder funds/Total assets (%)								
Name	1997	1998	1999	2000	2001	2002	2003	2004	average
FIRM A	41.4	62.5	57.6	97.7	99.5	97.5	94.9	93.1	80.5
FIRM D		99.7	99.7	96.4	99.6	99.5	99.7	99.7	99.2
FIRM E		100	100	100	100	100	100	100	100
FIRM F	16.0	13.4	85.2	68.7	80.2	69.7	69.1	57.9	57.5
FIRM G	61.3	65.8	71.6	69.8	66.7	67.5	65.6	66.5	66.8
FIRM C	55.2	61.3	60.0	59.1	68.5	75.9	76.8	80.3	67.1
FIRM I		100	100	100	100	22.8	21.3	20.2	66.3
FIRM B*									
FIRM K	26.0	21.0	27.1	31.0	39.8	37.2	20.2	6.7	26.1
FIRM H	95.0	95.0	95.0	95.0	96.6	98.6	99.3	99.3	96.7
FIRM J	99.5	94.4	92.3	100	100	100	100	2.1	86.0
Aggregated firm*	22.6	34.4	73.9	88.2	93.2	86.5	84.9	78.1	70.2

* Firm B is excluded because it reports negative shareholder funds.
Firms' names are fictitious

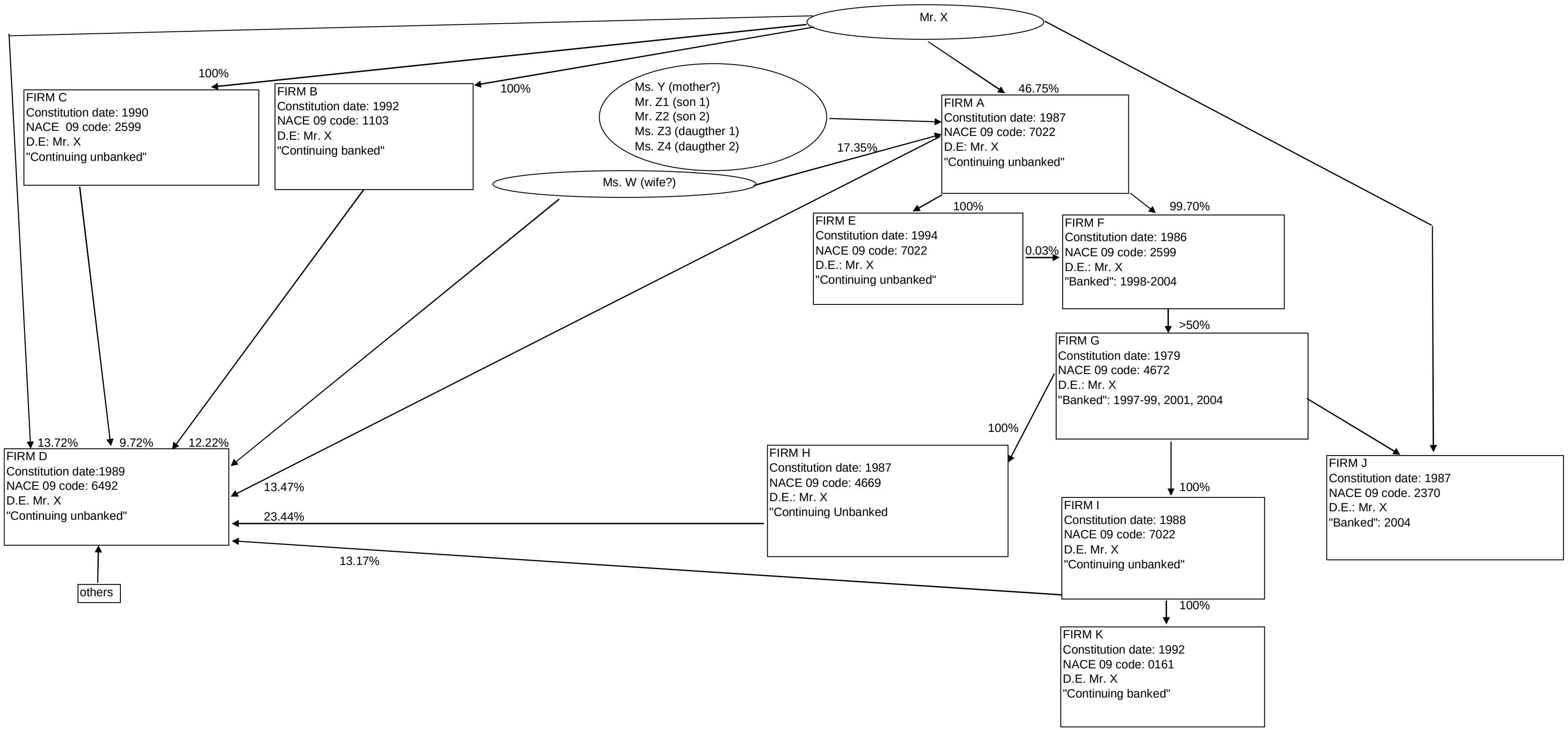
[73] The anomalously high ROA in year 2000 is explained by the very high financial revenues in that year.

Table A.6.2. Bank debt-to-total assets ratio (%)[1]

	Bank debt/Total assets (%)								
	1997	1998	1999	2000	2001	2002	2003	2004	average
FIRM A	0	0	0	0	0	0	0	0	0.0
FIRM D		0	0	0	0	0	0	0	0.0
FIRM E		0	0	0	0	0	0	0	0
FIRM F	0	0.1	0.2	2.1	0.6	8.5	16.5	10.0	4.7
FIRM G	3.9	7.0	3.0	0	2.1	0	0	3.2	2.4
FIRM C	0	0	0	0	0	0	0	0	0.0
FIRM I		0	0	0	0	0	0	0	0.0
FIRM B	20.6	31.9	39.1	39.2	14.8	14.1	10.7	10.7	22.6
FIRM K	22.9	20.8	66.5	50.0	39.8	29.6	23.0	42.9	36.9
FIRM H	0	0	0	0	0	0	0	0	0.0
FIRM J	0	0	0	0	0	0	0	43.6	5.4
Aggregated firm	0.5	0.7	1.5	2.0	1.0	2.9	4.8	4.5	2.2

[1]Information on banking debt has been drawn from CIR.
Firms' names are fictitious.

Chart A6: An example of family firms' structure:



Percentages associated to arrows indicate the direct ownership percentage of a shareholder (arrow origin) in a given firm (arrow end).
D.E: Director executive
Names are fictitious