**Family Firms, Bank Relationships and Financial Constraints: A Comprehensive Score Card**[1]

Alexander Karaivanov, Simon Fraser University
Jesús Saurina, Bank of Spain
Robert M. Townsend, MIT

May 2018

**Abstract**

We examine the effect of financial constraints on firm investment and cash flow. We combine data from the Spanish Mercantile Registry and the Bank of Spain Credit Registry to classify firms according to whether they are family-owned, not family-owned, or belong to a family-linked network of firms and according to their number of banking relations (with none, one, or several banks). Our empirical strategy is structural, based on a dynamic model solved numerically to generate the joint distribution of firm capital (size), investment and cash flow, both in cross-sections and in panel data. We consider three alternative financial settings: saving only, borrowing and lending, and moral hazard constrained state-contingent credit. We estimate each setting via maximum likelihood and compare across these financial regimes. Based on the estimated financial regime, we show that family firms, especially those belonging to networks based on ownership, are associated with a more flexible market or contract environment and are less financially constrained than non-family firms. This result survives stratifications of family and non-family firms by bank status, region, industry and time period. Family firms are better able to allocate funds and smooth investment across states of the world and over time, arguably done informally or using the cash flow generated at the level of the network. We also validate our structural approach by demonstrating that it performs well in traditional categories, by stratifying firms by size and age and find that smaller and younger firms are more constrained than larger and older firms.

JEL classification: C61, D82, D92, G21, G30
Keywords: financial constraints, family firms, bank lending, structural estimation and testing

[1]Research support from the 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 Espana or the Eurosystem. We thank V. Salas and K. Ueda for comments and suggestions on earlier versions. We also appreciate the excellent comments from conference audiences at the Far East Meeting of the Econometric Society, the European Economic Association, the Society of Economic Dynamics, and the Econometric Society World Congress. We are grateful to S. Ruano for her valuable contributions to an earlier and significantly different version of the paper and to J. Galan for his excellent research assistance with the current version.

# 1 Introduction

We focus on heterogeneity in financial constraints for firms, creating a comprehensive 'score card' from our findings. We integrate data and theory and analyze jointly firms' investment, cash flow and capital data from two large comprehensive databases on Spanish non-financial firms, the Spanish Mercantile Registry and the Bank of Spain Credit Registry (CIR). We then structurally estimate the underlying financial regime, among a range of alternatives, so as to be specific about (some of) the mechanisms determining the financial constraints, through the lens of theory, from the least to most constrained: moral hazard, borrowing/lending, to saving only. The data are stratified by a large number of cells, that is, by age, size, family ownership, number of banking relations, region, and industry. We also analyze the data over three different key sub-periods. To our knowledge, such a comprehensive score-card approach, combining a variety of structural models with a number of key stratifications of the data, is unique. A further advantage of our approach is that the models of financial constraints that we test are dynamic; that is, we explicitly model and pay attention to the firms' intertemporal incentives to save and invest, for example, constrained now versus constrained in the future.

Estimation results using both cross-sectional and panel-data show that family firms data are better fit by less constrained financial regime than the corresponding data for non-family firms. This points towards an advantage of family networks in accommodating investment needs to cash flows obtained in different firms which are parts of the network. By shifting funds from firms with less investment opportunities and more cash flow to others where the opposite holds, family networks may surmount financial constraints that otherwise could put downward pressure on firms' investment levels and smoothing capabilities. We show that this result is robust across firm age and size, time period, region, industry, and banking status strata. That is, while there are important differences in firm characteristics across these stratifications, once we control for that by estimating within the respective sub-samples, we still uncover the same major pattern regarding family vs. non-family firms.

Looking within the group of family firms, we find strong evidence that family-networked firms (that is, firms in a family-based network constructed by ownership shares) face less strict financing constraints compared to non-networked family firms ('pure family' firms, which are individually or family owned but not owners of other firms). In terms of the estimated financial regime and constraints, 'pure family' firms tend to fall in between the non-family and family-networked firms and are more similar to non-family firms than to family-networked firms.

We also find a clear pattern by firm age and size – estimating using data from larger and older firms yields a less constrained best-fitting financial regime compared to using corresponding data for small and young firms. These results are consistent with the previous findings from the empirical literature that larger and older firms tend to feature less sensitive investment / cash flow relationship compared to smaller and younger firms and thus validate our structural method.

Stratifying by firms' banking status, our findings depend on the firms' size and age. In the whole sample and among small and young firms, the estimated best-fitting financial regime for unbanked firms is more constrained than the best-fitting regime for single- and multi-banked firms. In contrast, we see a different pattern within the category of large and old firms – in the period 2001-2007 large and old unbanked firms appear less financially constrained than single-banked or multi-banked firms. A possible explanation is the different composition by industry and/or that these firms may have saved or financed their way out of constraints. We report extensively on this heterogeneity and our attempts to control for selection.

As noted at the outset, we use data from two large comprehensive databases on Spanish non-financial firms. First, we use the Spanish Mercantile Registry to obtain balance sheets, profit and loss accounts, information on financial/real flows (capital, investment, cash flow) and firm characteristics (ownership, age, etc.) Second, we use 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.[2] Since the minimum reporting threshold was low (6,000 Euros per loan) over our sample period, the CIR data are essentially a census of all banking loans granted to non-financial firms in Spain. In the estimation, as in the model, we focus on firms that continuously maintain the same banking status (unbanked, single-banked or multi-banked) in all years of data.[3]

We construct an indicator of whether a firm is family owned, not family owned or part of a family network defined by ownership shares obtained from the firm registry. We first identify all firms that are directly family owned (50% or larger share is held by an individual or family). We then add to this set of family firms, all other firms owned with 50% or larger share by the family firms in the initial set. We repeat the procedure until no new firms are added at further iterations. The end result is all firms controlled by an individual or family, directly or indirectly. We further sub-divide family firms into two types: 'pure family' firms (directly owned by individuals or families and which do not own other firms) and 'family-networked' firms (part of a network of firms constructed by ownership). The rationale is to distinguish different family firms, e.g., a 'mom-and-pop store' vs. a company within a large business group.

In more detail, we analyze the investment behavior of Spanish firms taking into account their age, size, bank relationships, and ownership structure. We also stratify the firms by region (Madrid, the Mediterranean and the Ebro river valley vs. the rest of the country) and by industry (construction and real estate vs. all other non-financial sectors). We analyze firm investment in a country where business funding is overwhelmingly dominated by banks and in a time period that corresponds to the longest continuous expansion of the Spanish economy under a lending boom that fuelled firms' investment expansion. We analyze three sub-periods in a balanced manner: 1997-2000, which corresponds to the investment expansion after the 1993 recession in the wake of entering the Eurozone; 2001-2003, an economic slowdown as a result of recession in core Euro zone countries; and finally 2004-2007, the years right before the financial crisis and also a time when lending and liquidity was at its peak. While the 1997–2007 period we study corresponds to different growth periods of the Spanish economy, we do not find a strong pattern in terms of the best-fitting models of financial constraints over time, however, we do see more instances of the least constrained (moral hazard) regime estimated as best-fitting in the 2004-07 data, the period of the credit ramp-up before the financial crisis.

Our definition of financial constraints is about how damaging, or alternatively how flexible, is the financing/information regime cell in which a firm resides, based on the estimation of alternative dynamic models. The financial regimes range from saving only to unrestricted borrowing/lending with an institution or market to moral hazard constrained insurance and credit arrangement. We tabulate the placement of firms into these cells based on key characteristics emphasized in the literature: size and age, bank finance (no, single, or multiple bank lenders), family status (non-family, family-owned, in a family-linked network determined by ownership), region, industry and time period. We further disaggregate the data by intersecting these categories, for example, small and young

[2] Jiménez et al. (2012, 2014, 2016) have used extensively both databases to analyze monetary policy and financial stability. Here we focus on the real side of the economy by investigating non-financial firms' investment.

[3] We implicityly assume that multi-bank lending delivers an outcome as if a firm were dealing with a unified sector, as when banks can coordinate, but this is a limitation of the model.

single-banked firms that are family connected vs. small and young single-banked with no family connections, etc. The bottom line is a score card or indicator for which group of firms is more severely financially constrained and which less so.

A large literature on firm investment at the micro level studies the role of financial constraints.[4] Some authors use models of information or incentive problems in capital markets to motivate the role and origin of endogenous financial constraints.[5] Other authors assume exogenously incomplete credit markets, for example borrowing and saving in a single asset. In both cases the ultimate result is limited access to external finance – firms have to restrict investment when internal cash is insufficient to invest at the first-best level.[6] Compared to previous approaches relying on more reduced-form analysis, including using panel data,[7] the main advantage of our structural approach is that we can assess not only whether financing constraints affecting firm investment are absent or present (perfect vs. imperfect credit markets), but also determine the most likely nature of the financial constraints, within the taxonomy of the structural models we compute, estimate and statistically test. Here we feature three prototypical models of exogenously or endogenously incomplete markets but, as shown in Karaivanov and Townsend (2014), other types of financial constraints can be easily added. For example, a complete markets setting was considered in earlier drafts but rejected by the data. We explicitly model the dynamics and interaction of present vs. future constraints as firm size and credit conditions (proxied by state variables in the models), evolve endogenously over time within each financial regime. For example, a firm could have high cash flow in the current period but anticipate constraints in future periods and so its current investment may not respond much to current cash flow. Alternatively, the firm may carry past debt (in the borrowing/lending model), or have low promised utility due to having been net transfer recipient in the past (in the moral hazard model), which would affect its current investment. Our approach takes these dynamic interactions into account, both with respect to their implications in the cross-section and for panel data.

However, we also offer an important caveat about our structural approach at the outset: no model or set of models is sufficiently comprehensive to incorporate all possible factors affecting firms' investment decisions and constraints. We do our best to account for heterogeneity in observables by using the different data sub-samples by size and age, region, industry, etc. We also do our best to be clear about the limitations of our approach, but not exclusive to it, especially the important issue of selection on unobservables, though heterogeneity in unobservables is admitted in the models we consider.

First, to be consistent with the theory and control for the possible endogeneity of bank relationships, in constructing the sample used in the structural estimation, we focus only on firms that remain continuously in the same bank status category for all years in the data. The firms' ownership status (family or non-family) is also treated as time invariant as we have only one measure for it in the data. The underlying assumption is that constant banking and ownership status proxy for the firms staying in the same financial regime over time – the dynamic programming problem implies that the firms would behave approximately as if they were set to be in that category

[4]Schiantarelli (1996), Hubbard (1998) and Strebulaev and Whited (2012) provide extensive reviews.

[5]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 designed contracts when profit outcomes can only be observed at a cost. On costly state verification see also Townsend (1979) and Diamond (1984).

[6]The cost gap between external and internal funds could be explained by information frictions but could also be due to taxes or other transaction costs.

[7]See, among others, Bond and Meghir (1994) and Bond et al (2003).

forever, as if at the end of the sample they are still not near making a transition. We discuss this approximation in the theory section. This assumption, however, cannot be directly tested. Moreover, when we estimate using panel data for our constructed categories, we need a balanced panel in order to compute the joint likelihood of investment and cash flow over time. Nonetheless, we also perform robustness checks with all available data.

Second, firms within a given banking status or family category could be substantially different from those in another category for other reasons. We do incorporate endogenous, unobserved-to-the-econometrician variables within the models – unobserved 'promised utility' (present discounted value of profits) in the moral hazard model and unobserved debt/savings in the borrowing model. Any other remaining heterogeneity is not yet incorporated into the models. We did estimation runs in which we removed firm fixed effects from the capital stock, investment and cash flow data. Doing so, however, removed a substantial part of the data variation and the Vuong comparison test was unable to pick up statistically significant differences in the financial constraints between the firm categories – for example in all stratifications by bank or family status that we tried the borrowing and lending model was best-fitting. We perform and report on various other robustness checks, with respect to both the data sample (all firms, all continuing banking status firms, placebo family status, switching banking status firms) and the model (allowing for risk aversion or adjustment costs).

**Literature review**

The empirical literature has documented that, in addition to the fundamentals determining investment dynamics such as expected future profitability and the user cost of capital, firms' internal funds influence 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 et al. (1988), a widely used empirical strategy consists of separating 'constrained' and 'unconstrained' firms according to a priori assumptions on the likelihood that they are subject to information or incentive problems, then testing whether the neoclassical investment equation derived under the assumption of perfect capital markets with adjustment costs describes the behavior of unconstrained firms but not that of constrained firms.[8] Typical firm characteristics used as a priori proxies for the importance of capital market imperfections are firm's size, age, and dividend policy.[9] A few papers stratify by relationship with industrial or financial groups, and this is related to our findings. In our structural approach we try to avoid some of the back-and-forth in the debates concerning the mapping between reduced form empirical work, on the one hand, and theory on the other, though in our view these debates have been productive (e.g., Kaplan and Zingales,1997 and 2000 vs. Fazzari et al., 1988 and 2000).[10]

We study the importance of family ownership, on its own and in conjunction with age, size, banking status,

[8]There are two main empirical approaches, pioneered by Fazzari et al. (1988) and Bond and Meghir (1994) respectively. The former is based on estimating the Q-model of investment, extended by 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 estimates an investment Euler equation which should hold for unconstrained firms but be mis-specified for constrained firms. Both approaches assume perfect competition, constant returns to scale and quadratic adjustment costs.

[9]Schiantarelli (1996) provides an exhaustive list and discussion.

[10]As argued by Kaplan and Zingales (1997, 2000), a firm is either constrained or unconstrained, a binary variable, and thus the degree of cash flow sensitivity does not necessarily indicate the source or severity of financial constraints. Gomes (2001), Cooper and Ejarque (2003) and others have further challenged the Fazzari et al. (1988) interpretation by demonstrating that financing constraints are neither necessary nor sufficient for cash flow effects. 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). On the other hand, Schmid (2009) shows that, in a model with incomplete markets due to limited commitment, the cash flow effect is significant and so suitably modelled financial frictions could rationalize the evidence on cash flow sensitivity.

region, industry and time period, for the degree and nature of financial constraints faced by firms. There is a literature on the positive and negative attributes of family-owned firms (Caspar et al., 2010). Khanna an Yafeh (2007) characterize business groups, with particular attention to family groups, as 'paragons' or 'parasites'. Bertrand and Schoar (2006) review world-wide evidence on family controlled businesses, contrasting arguments for the higher efficiency of family firms vs. cultural theories claiming the contrary. Specific examples include risk sharing in family networks, as in Kinnan and Townsend (2012) for Thai SME's and Samphantharak (2003) for larger conglomerates, versus the literature on 'tunnelling' as in Bertrand et al. (2002) and others. Papers on family-tied firms in Europe include Cronqvist and Nilsson (2003) for Sweden, Sraer and Thesmar (2004) for France or Maury (2006) for a sample of Western European countries (see also Morck, 2005 for an exhaustive historical review).

We hypothesize, and confirm in our empirical results, that ownership connections between family firms and the resulting network they form may increase these firms' financing possibilities and may allow such networked firms, including unbanked, to behave as if they face less strict financial constraints compared to family firms that are not part of a network or compared to non-family firms. Our results on family-owned and family-networked firms are in line with the literature that shows the importance, both from a theoretical and an empirical point of view, of internal financial markets in allocating resources while competing with external funding sources. In Spain higher levels of finance beyond bank borrowing come not from bonds but from trade credit and the Spanish analogue to chaebols or zaibatsu business groups (Hoshi et al., 1990). Pindado and De La Torre (2011) study whether family control alleviates or exacerbates investment cash flow sensitivity in the Euro zone. They find that family-controlled firms have lower sensitivity and this is mainly attributable to family firms with no deviations between cash flow and voting rights and to family firms in which family members hold managerial positions. Consistent with our results, obtained using a completely different structural methodology, the authors argue that family control seems to mitigate investment inefficiencies that derive from capital market imperfections.[11]

Related is the large literature on bank finance. Banks are thought to have a comparative advantage in producing information about borrowing firms (Diamond, 1984; Bond, 2004). To the extent that bank relationships mitigate asymmetric information problems and facilitate monitoring, the incidence of financial constraints on firms' investment and its sensitivity to cash flow can differ across firms stratified by the strength of their bank relationships.[12] The empirical research evaluating the costs and benefits of bank relationships is extensive, including papers studying differences in firms' investment behavior. In most cases, closer bank relationships are associated with lower sensitivity of investment to cash flow (Elston, 1996 for Germany; García-Marco and Ocaña, 1999 for Spain; and Houston and James, 2001 for the USA). However, some authors find no significant effect of bank relationships on cash flow sensitivity (Fuss and Vermeulen, 2006 for Belgium) or a positive effect (Fohlin, 1998, for Germany).[13] For Spain, Carbo-Valverde et al. (2008) find that the investment of SMEs is sensitive to

[11]Andres (2011) finds a similar result using German data – when compared to firms of similar size and dividend payout ratio, the investment outlays of founding family owned firms are consistently less sensitive to internal cash flow (see also Anderson et al., 2003 for related evidence on US founding family owned firms having lower cost of debt financing). Crespi and Martin-Oliver (2015) further document that during crisis periods, family firms are less subject to credit constraints than non-family firms and thus the impact of lack of credit on their capital structure is less severe. Lopez-Gracia and Andujar-Sanchez (2015) find that growth opportunities, financial distress costs, and internal resources are the main factors that differentiate family vs. non-family businesses in Spain.

[12]The strength of bank relationships has been defined in the literature as: the relationship's length (Petersen and Rajan, 1994; Berger and Udell, 1995 and Degryse and Van Cayseele, 2000); scope, that is, 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;) or the number of banks (Petersen and Rajan, 1994; Houston and James, 2001 and Fuss and Vermeulen, 2006).

[13]Other papers investigate differences in the role of bank relationships in alleviating financial constraints at different stages of the

bank loans for unconstrained firms but not for constrained firms (defined as firms with desired borrowing larger than the amount of credit provided by banks) and that trade credit predicts investment but only for constrained firms.[14] In addition, unconstrained firms use bank loans to finance trade credit provided to other firms.[15]

We use our Bank of Spain data on bank loans and categorize non-financial Spanish firms into continuously unbanked, single-banked and multi-banked. At one extreme is the literature on financial access which finds that, without formal credit access (unbanked), businesses can be quite constrained, hence with high marginal returns although not always accounting for risk premia (De Mel et al., 2008; Evenson and Gollin, 2003; Udry and Anagol, 2006; Samphantharak and Townsend, 2016). Having access to a single bank can be thought a priori to be helpful – there is a large literature on relationship banking based on the premise that information about borrowers is extremely important in the lending process, especially for small firms which are often considered informationally opaque.[16] The literature on multiple bank access draws mixed conclusions. On the one hand, financial access from two or more lenders could mean more available credit and hence less binding financial constraints. On the other hand, allowing borrowing from multiple sources may lead to enforcement or adverse selection problems which relates to the literature on coordination failures without credit registries and contingencies in loan contracts (Bizer and DeMarzo, 1992; Green and Liu, 2015). More generally, there is both theoretical and empirical research which finds that competition among banks can be either good or bad depending on the setting.[17]

We also follow the literature and stratify our sample by firm age and size. This is the approach of Fazzari et al. (1988), with the idea that the smaller and/or younger is a firm, the more likely it is to be constrained. There are also life-cycle firm finance models in the theoretical literature which reach the same conclusion (Clementi and Hopenhayn, 2006 or Albuquerque and Hopenhayn, 2004, featuring either moral hazard or limited enforcement). A large development literature on financing constraints makes current net worth a constraint on credit (Banerjee and Duflo, 2005; Gine and Townsend, 2004; Jeong and Townsend, 2007) and part of this literature focuses on persistence and how businesses save over time to overcome constraints (Buera and Shin, 2013; Moll, 2014; Banerjee and Moll, 2010; Mestieri et al., 2016).

Regarding regional effects, Fabbri (2010) analyzes the effect of judicial efficiency on firm financing in Spain. The author uses balance sheet data on 1,700 Spanish manufacturing firms and regional data on judicial trials (average length and percent completion within a year). Fabbri finds that in districts with lower judicial efficiency bank financing is costlier and firms are on average smaller but there is no significant effect of judicial efficiency on the firms' leverage ratio.

Finally, in terms of method, the computational and empirical techniques we use are based on Karaivanov and Townsend (2014) where we estimate and statistically test dynamic models of consumption, income and investment

credit cycle. For example, Vickery (2005) shows that the benefits of bank relationships are higher when credit conditions are poor or deteriorating. Direct evidence from matched data suggests that borrowers are unable to smooth bank liquidity shocks completely (Khwaja and Mian, 2008; Jiménez, et al, 2012; Schnabl, 2010). Some authors also find that bank capital shocks have real consequences in the form of decreased investment by borrowers (Gibson, 1995; Kang and Stulz, 2000; Gan, 2007; Chava and Purnanandam, 2009).

[14] Although not focusing on bank relationships, the influence of financial constraints on the investment decisions of Spanish firms is also analysed in Alonso-Borrego (1994), Estrada and Vallés (1998) and Hernando and Tiorno (2002).

[15] Fohlin and Iturriaga (2010) study the relationship between investment-cash flow sensitivity (ICFS), bank relationships and firm ownership structure. They find that bank relationships via equity or debt have little effect on Spanish firms' ICFS. In contrast, firms with high ownership concentration exhibit significantly lower sensitivity. D'Espallier et al. (2011) use a Bayesian approach to study ICFS among 90 listed Spanish firms and provide evidence that high-ICFS firms have higher financing needs while faced with fewer available external financing sources.

[16] For comprehensive reviews on relationship banking see Boot (2000) or Elyasiani and Goldberg (2004).

[17] Petersen and Rajan (1995), Cetorelli (2001), Alessandrini et al., (2009), Degryse and Ongena (2005) and Zhorin and Townsend (2014).

with data on households running small businesses in rural and (semi-)urban Thailand. In the rural sample, we find that a saving only or borrowing in a single asset regime provide best fit using joint data on consumption, assets, investment, and income. In contrast, a moral hazard model fits best the data in the urban sample. In addition to the different setting, here we focus on firm investment and cash flow and on the role of bank relationships and firm ownership. In Karaivanov and Townsend (2014) we also found that family or loan/gift networks help households achieve better consumption smoothing; similarly here, but using completely different data, we find that family-networked firms exhibit investment/cash flow pattern corresponding to less severe financial constraints both in the cross-section and over time. More generally, our approach here relates to other papers that use data to try to tell apart different models of financial constraints, although not always in a fully structural way (Ligon, 1998; Paulson and Townsend, 2004; Paulson et al., 2006; Ahlin and Townsend, 2007; Dubois et al., 2008; Karaivanov, 2012; Kinnan, 2014).

## 2 Theory

We model a firm's payoff as $u\,(c, z) = c - \nu(z)$ where $c \geq 0$ is dividends to the owner, analogous to 'consumption', and $z$ is 'effort' in management and production. The cost of effort $\nu(z)$ is strictly increasing. Preserving separability in $c$ and $z$, the function $u\,(c, z)$ can be generalized to accommodate risk aversion (see the robustness section).

Following the firm investment literature (e.g., Bond and Meghir, 1994), the firm's production function maps current costlessly adjustable inputs, $z$ and capital, $k$ into net profits or cash flow, $q(k, z, \theta)$ where $\theta$ is a firm-specific shock. Cash flow $q(k, z, \theta)$ equals revenue net of expenses for hired labor and other costs which we do not write explicitly. This specification allows the shocks $\theta$ to be autocorrelated, which is important in the cash flow versus productivity debate in the literature. Here we focus on the case in which, *conditional on $k$ and $z$*, the shocks $\theta$ are i.i.d. over time and across firms.[18] However, since the input choices $k$ and $z$ are endogenous and depend on the state variable values over time (capital, promised utility or debt/savings) and since capital does not depreciate fully, our model yields endogenous autocorrelation in cash flow $q$. This serial correlation is one of the dimensions of the joint distribution of the data variables, $k, I$ and $q$ that we use in the MLE to estimate the models with panel data.

Suppose the shock $\theta$ can take a finite number of values, $\{\theta_1, ..., \theta_m\}$ and let $q_i$ be cash flow (net profits) in state of the world $\theta_i$. Call $Q \equiv \{q_i\}_{i=1}^m$ the set of all possible cash flow levels. The probability of realizing a particular $q_i$ depends on the firm's capital $k$ and the effort input $z$ used in production. To capture the stochasticity of $q(k, z, \theta)$ and its dependence on capital and effort, we denote by $P(\tilde{q}|k, z)$, for any $\tilde{q} \in Q$, the probability of a firm having cash flow $\tilde{q}$, conditional on its capital stock, $k$ and effort level, $z$.[19]

Firm capital depreciates at rate $\delta \in (0, 1)$ and follows the law of motion,

$$k_t = (1 - \delta)k_{t-1} + I_t,$$

where $I_t$ denotes period-$t$ investment. We do not constrain investment to be positive, that is, capital accumulation

[18] Since effort $z$ is unobservable, we are not able to distinuish in the data possible persistence in $\theta$ vs. persistence in $z$.

[19] We discuss how $P$ is constructed from the data, with additional parametric assumptions to account for the unobservable $z$, in Appendix A2. We find that the matrix P differs very little across the time periods we study.

is reversible.[20] As in Bond and Meghir (1994), we assume that current investment $I_t$ adds immediately to the capital $k_t$ used in the current-period production.[21] We can also allow for cost $g(k, I) \geq 0$ of adjusting the capital stock, with $g(k, I)$ convex in investment $I$ and subtracted from the firm's net profits $q$. We perform robustness estimation runs with adjustment costs (see Section 4).

The firm discounts future profits with factor $\beta \in (0, 1)$. In what follows we set $\beta$ equal to the inverse of the borrowing and lending market rate, $\beta = 1/R$, where $R = 1 + r$ is the gross interest rate, but more generally, we can allow $\beta$ and $1/R$ to differ. We can also easily incorporate two interest rates, $R^B$ for borrowing and $R^S$ for saving, with $R^B \geq R^S$.

## 2.1 Financial regimes

We consider three alternative financial regimes in which a firm may operate. The regimes are interpreted as representing financial market settings which differ in their level of sophistication or credit constraints. First, a firm could be in what we call a *saving only* (S) regime. Such firm is assumed to have no access to credit but it can save at a fixed gross rate of return $R$. Second, a firm could be in what we call a *borrowing and lending* (B) regime. Such firm has access to a financial market allowing it to borrow or save in a non-contingent asset with fixed gross interest rate $R$. Third, a firm could operate in a more complex financial regime with contingent financial assets and subject to endogenous financial constraints because of an information problem (unobservable actions) – a *moral hazard* (MH) regime.[22]

**Saving only (S) / Borrowing and lending (B)**

In these financial regimes a firm can save (or also borrow, in regime B) at an exogenously given gross market interest rate $R$. This can be thought of the firm trading in a single non-contingent asset with fixed return, similarly to the permanent income hypothesis setting (Bewley, 1977). There is no possibility of default or contingencies of any kind. Given beginning-of-period capital stock $k$ and debt $b$, the firm's problem in this exogenously incomplete markets setting can be written as:

$$
\begin{aligned}
\Pi(k, b) \quad = \quad & \max_{\{k', b'(q), c(q), z\}} \sum_{q \in Q} P(q|k', z)[u(c(q), z) + \beta \Pi(k', b'(q))] \\
& \text{s.t. } c(q) + k' - (1 - \delta)k + Rb - b'(q) = q - g(k, k' - (1 - \delta)k) \text{ for all } q \in Q
\end{aligned}
$$

where $b'(q)$ denotes next period's debt or savings[23] and where, as in Bond and Meghir's (1994), the state variable $k$ is beginning-of-period capital stock while $k'$ denotes the post-investment capital used in current production. Current period investment, $I$ equals $k' - (1 - \delta)k$. The budget/feasibility constraint holds for each possible cash flow realization $q$. Dividends $c(q)$ are contingent on the cash flow realization.

[20]In the data, investment defined as $k_t - (1 - \delta)k_{t-1}$ is sometimes negative; for example, this is true for 6.7% of all observations in the 2004-07 panel data.

[21]In Bond and Meghir (1994) investment is subtracted as cost in the net profit function. Note also that in our paper cash flow, $q$ is already net of wages and other costs. In Bond and Meghir (1994) these non-capital input decisions are explicit.

[22]We have chosen the moral hazard regime as a prototypical setting with endogenously incomplete financial markets. Other models (e.g., hidden income, limited commitment) can be used within our empirical method (see Karaivanov and Townsend, 2014) but we are constrained here computationally, as each single estimation run takes up to 1 day and we have more than 200 runs across the different data stratifications we consider.

[23]Positive values of $b$ are interpreted as debt, negative values are interpreted as savings.

In the saving only (S) regime we set the upper bound on debt to $\bar{b} = 0$ (that is, $b \leq 0$ and $b'(q) \leq 0$ for all $q$). In the borrowing regime (B) the assumption of no default implies an endogenous borrowing constraint $\bar{b}$ (if debt exceeds $\bar{b}$ the problem has no solution).

**Moral hazard constrained credit (MH)**

To model a more complex financial setting which allows for risk-contingent premia, transfers and debt, we write the firm owners' payoff as $u(c, z) + \beta w'$, where

$$c = \tau + (1 - \delta)k - k' - g(k, I)$$

is current dividends, $k$ is the beginning-of-period capital stock and $k'$ is the post-investment capital used in production. The variable $w'$ is next period's 'promised utility', that is, discounted expected future payoff. In the risk-neutrality case it equals the discounted value of all expected dividends from next period onward. The firm is assumed to have access to a risk-neutral lender with discount factor $1/R$. The lender, or collection of lenders, is modelled as a principal maximizing discounted expected returns when facing a long-term contract with a firm with current promised utility $w$ and initial capital $k$. Both the lender and the firm have full commitment.

In our formulation, it is as if all net profits $q$ go to the lender but some part is then returned to the firm via the transfer $\tau(q)$ contingent on the realization of net profits $q$. The firm bears the costs of changes to its capital stock, that is, it pays for investment (and adjustment costs) but, assuming full observability, its capital stock and investment are effectively under the control of the lender. In contrast, the effort $z$ is unobserved by the lender and so the firm must be given incentives to perform – a moral hazard problem arises.

To write the contracting problem in this setting, call $V(k, w)$ the risk-neutral lender's (or collection of lenders) profits from facing a firm with beginning-of-period capital, $k$ and current promised utility, $w$ belonging to the set $W$ (to be determined below). The lender solves:

$$V(k, w) = \max_{\{k', w'(q), \tau(q), z\}} \sum_{q \in Q} P(q|k', z)[q - \tau(q) + (1/R)V(k', w'(q))] \tag{1}$$

subject to two constraints. First, the contract must satisfy a *promise keeping* constraint stating that, in expectation, the firm must receive discounted payoff (the l.h.s.) equal to the promise $w$,

$$\sum_{q \in Q} P(q|k', z)[u(c(q), z) + \beta w'(q)] = w \quad \text{(PK)}$$

where $c(q) \equiv \tau(q) + (1 - \delta)k - k' - g(k, k' - (1 - \delta)k)$.

Second, because of the moral hazard problem, the lender faces the following incentive compatibility constraint,[24]

$$z = \arg\max_{\tilde{z}} \sum_{q \in Q} P(q|k', \tilde{z})[u(c(q), \tilde{z}) + \beta w'(q)] \quad \text{(ICC)}$$

The moral hazard setting we consider assumes full commitment and exclusivity. It is known from the literature that if the exclusivity assumption is relaxed or the agent can engage in hidden trades, the constrained efficient

[24] The setting of perfect credit markets (full information), in which the borrower's effort $z$ is contractible is also easy to model and compute by maximizing (1) subject to the participation constraint (PK) without imposing the incentive constraint (ICC).

contract can be equivalent to a simple debt contract (Allen, 1985; Cole and Kocherlakota, 2001; Kirpalani, 2016; Ales and Maziero, 2016). While we acknowledge that the assumption of (non-)exclusivity is important in terms of how the financial constraints are interpreted, we do feature a range of financial regimes, including borrowing and lending, and let the data pick the best-fitting regime.[25]

In our empirical analysis in Section 4 the firm's capital stock $k$, investment $I$ and cash flow $q$ are observables, subject to measurement error with distribution we estimate, while effort $z$ is an unobservable choice variable. Firm debt or savings, $b$ in the S and B regimes and promised utility, $w$ in the MH regime are treated as sources of unobservable heterogeneity with a parametric distribution that we estimate (see Appendix A for details). The production function $P(q|k, z)$ is derived from the cash flow and capital data, via a histogram function (see Appendix A).

The dynamic models of firm financial constraints presented above assume that the firms' financial access status does not change over time, only the state variables $k$ and $b$ (or $w$) evolve. Theoretically, this approximates firms that are in a "steady state" with regards to the financial environment and the constraints they face, that is they are (possibly infinitely) far away in time from a potential regime transition. In principle, our model can be extended along the lines of Townsend and Ueda (2010) to allow endogenous transitions between financial regimes over time but we abstract from this here.[26]

In sum, we model three dynamic financial regimes ranging from the most constrained (saving only) to the least constrained (moral hazard) in terms of availability of credit and ability to smooth investment over time and across states. We estimate and test these three regimes based on their implications for firm investment and its sensitivity to cash flow using maximum likelihood, as explained in Section 4.

## 2.2 Example

To illustrate the differences between the three financial regimes in their implications for firm assets, $k$ investment, $I$ and dividends, $c$ for any given realized net profits /cash flow $q$, we solve a one-period version of our model. To keep things as simple and tractable as possible, assume that there are only two possible cash flow levels, $q_L$ and $q_H$ with $q_H > q_L > 0$. The probability that $q_H$ is realized is $p(k', z)$ where the function $p$ is strictly increasing in both arguments and has strictly decreasing partial derivatives satisfying Inada conditions. The firms are risk neutral with payoff function $u(c, z) = c - \nu(z)$ where $v$ is strictly increasing and convex. There are no capital adjustment costs.

**Saving only / Borrowing and lending (S / B)**

In the one-period setting the firm cannot save or borrow and hence the S and B regimes coincide. Given initial capital $k_0$, the firm's problem is to choose capital $k$ (equivalently, choose investment $I = k - (1 - \delta)k_0$) and effort

[25] Also, there are reliable credit registries in Spain and each lender knows how much each firm has borrowed from others.

[26] Townsend and Ueda (2010) consider autarky (agents exposed to idiosyncratic risk) and a financial sector providing full insurance to idiosyncratic risk as two extreme settings, with costly entry into the financial sector. They find that the optimal saving behavior in autarky is approximated the better the more distant is entry into the financial sector.

$z$ to use in production, as well as dividends $c$ in each state.

$$\max_{c_L, c_H, z, k} \; p(z,k)c_H + (1-p(z,k))c_L - v(z) \tag{2}$$
$$\text{s.t. } \; c_i = q_i + (1-\delta)k_0 - k \quad \text{for } i = L, H$$
$$\text{and } c_i \geq 0 \quad \text{for } i = L, H$$

The FOCs with respect to $k$ and $z$ at an interior solution are:

$$\frac{\partial p(z,k)}{\partial k}(q_H - q_L) = 1 \;\text{ and }\; \frac{\partial p(z,k)}{\partial z}(q_H - q_L) = v'(z), \tag{3}$$

after substituting for $c_i$ from the budget constraint. Call $k^*$ and $z^*$ the values solving the above two equations. Note that $k^*$ and $z^*$ are the same as the first-best capital and effort levels that an unconstrained firm would choose (risk neutrality is key for this).[27]

Suppose, given the Inada conditions and the assumed properties of $p(k,z)$ and $v(z)$, the FOC with respect to $z$ always has an interior solution. Then, the solution to problem (2), call it $(c_i^B, z^B, k^B)$, looks as follows. If $q_L$ and $k_0$ are sufficiently large so that the non-negativity constraint on dividends in the low state $c_L$ does not bind at $k^*$, that is, $q_L + (1-\delta)k_0 - k^* > 0$ then the firm chooses $k^B$ and $z^B$ at the unconstrained values $k^*$ and $z^*$ which implies $c_i^B = q_i + (1-\delta)k_0 - k^*$. If, however, $q_L$ and $k_0$ are such that the non-negativity constraint on $c_L$ binds (corner solution), then the solution to problem (2) has instead $c_L^B = 0$, $k^B = q_L + (1-\delta)k_0$, $c_H^B = q_H - q_L$ and $z^B$ solving the effort FOC evaluated at $k = k^B$. The firm would like to invest more but it is constrained.

**Moral hazard (MH)**

The one-period version of the firm's problem in the moral hazard regime can be written as follows, assuming initial promise $w_0$ that delivers zero profits to the lender,

$$\max_{c_L, c_H, z, k} \; p(z,k)c_H + (1-p(z,k))c_L - v(z)$$
$$\text{s.t. } p(z,k)c_H + (1-p(z,k))c_L + k = p(z,k)q_H + (1-p(z,k))q_L + (1-\delta)k_0 \quad \text{(ZP)}$$
$$\frac{\partial p(z,k)}{\partial z}(c_H - c_L) = v'(z) \quad \text{(ICC)}$$

where we used the "first-order approach" to replace the firm's incentive constraint with respect to the action $z$ with its first order condition (ICC). It can be shown that, with two output levels and $p(z,k)$ strictly concave in $z$, the first order approach is valid (see Karaivanov, 2012 for a proof).

[27] An unconstrained firm solves:

$$\max_{k, z, c_H, c_L} \; p(z,k)c_H + (1-p(z,k))c_L - v(z)$$
$$\text{s.t. } p(z,k)c_H + (1-p(z,k))c_L + k = p(z,k)q_H + (1-p(z,k))q_L + (1-\delta)k_0$$

Substituting for $p(z,k)c_H + (1-p(z,k))c_L$ from the resource constraint yields the same FOCs with respect to $k$ and $z$ as in (3).

Substituting from the zero-profits constraint (ZP) into the objective, the problem simplifies to

$$\max_{c_H, c_L, k, z} \; p(z,k)(q_H - q_L) - k - v(z) \tag{4}$$
$$\text{s.t. (ICC)}$$

Note first that, if we ignore (ICC) for the moment and assume interior solution, we would obtain the exact same FOCs with respect to $k, z$ as in the S/B setting. Therefore, choosing $k^{MH} = k^*$ and $z^{MH} = z^*$ and setting $c_H^{MH} - c_L^{MH} = q_H - q_L$ would also satisfy constraint (ICC) and hence solve problem (4). From (ZP) we then obtain that $c_i^{MH} = c_i^B$. Intuitively, at an interior solution, when the first-best effort and capital stock $z^*$ and $k^*$ are feasible, a *risk neutral* firm would choose them and the incentive constraint would be automatically satisfied (there is no trade-off between insurance and efficiency).

However, suppose that $q_L$ and $k_0$ are such that the non-negativity constraint on $c_L$ binds at $k^*$. In the S/B setting above, this required a constrained choice for capital $k^B = q_L + (1-\delta)k_0 < k^*$. In the moral hazard regime, the firm can do better and choose a higher level of investment, $k^{MH} \in (k^B, k^*)$ by "borrowing" from the high state at the expense of reducing $c_H$. The reason is that, unlike in the S or B regimes, only the cross-state constraint (ZP) must be satisfied. The downside of the reduction in $c_H$ is (all else equal) reduced effort relative to the first best. The exact outcome can be worked out numerically from this trade-off.

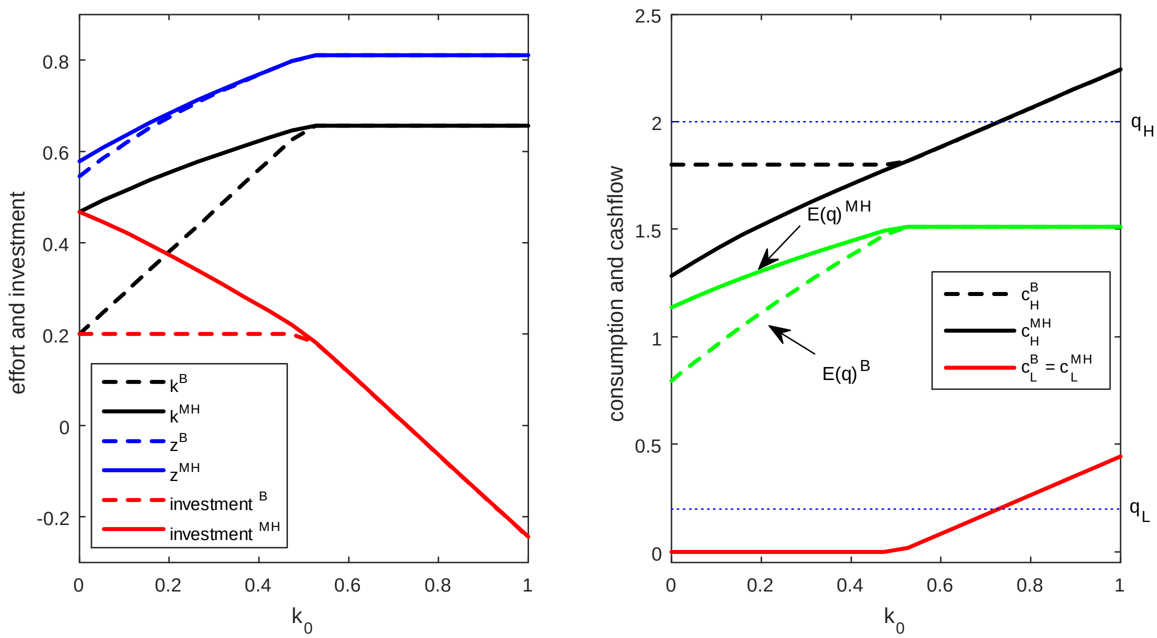
Figure 1 illustrates the solutions for capital, effort, investment ($I = k - (1+\delta)k_0$) and dividends $c_i$ in each state, as we vary initial capital $k_0$. The computed example uses $p(z,k) = \sqrt{kz}$, $u(c,z) = c - \frac{z^2}{2}$ and $q_H = 2$, $q_L = .2$, $\delta = .1$. For large $k_0$ (above .5 approximately) the S/B and MH solutions coincide. For these parameters the first-best input choices, $k^*$ and $z^*$ are feasible and incentive compatible. In contrast, for low values of $k_0$ the S/B and MH solutions differ in their implied investment, effort, expected cash flow, and dividends. Specifically, in the moral hazard constrained setting firms invest more on average and achieve higher expected cash flow.

**Discussion**

The one-period example shows that the implications of the moral hazard and saving/borrowing regimes regarding firms' investment and cash flow differ when the moral hazard problem has a "bite", that is when the first-best input levels $k^*$ and $z^*$ cannot be implemented without violating the non-negativity constraints on dividends $c$. The reason for the difference is that the one-period saving or borrowing setting is essentially financial autarky – the firm bears the full brunt of cash flow shocks while, in contrast, in the MH regime it is partially insured against these shocks as it can use the financial intermediary to transfer resources across states of the world.

Intuitively, under risk neutrality the moral hazard problem leads to inefficient choices $k$ and $z$ when the lower bound on consumption binds in the low cash flow state ($c_L = 0$). That is, the financial intermediary principal cannot 'punish' the agent sufficiently in case of low cash flow (raise the spread $c_H - c_L$) to induce the efficient level of effort. This parallels the more familiar moral hazard setting with risk aversion where the ability of the principal to provide incentives for supplying effort via raising the utility spread $u(c_H) - u(c_L)$ (see ICC) is limited by the need to provide insurance (that is, bringing $c_H$ and $c_L$ closer to each other). Figure 5 in the Appendix illustrates the model solution for the case of risk-averse firms with payoff function $u(c) - v(z)$, where $u$ is CRRA with coefficient of relative risk aversion equal to .1. The S/B and MH regimes yield different solutions for any initial capital $k_0$ as the moral hazard regime is always able to provide more smoothing.

Figure 1: Computed Solutions – Saving Only / Borrowing vs. Moral Hazard




**Multiple periods**

In a multi-period setting the intuition is similar to that in the one-period example, however, differences across all three financial regimes (S, B and MH) appear when the first-best is infeasible. For example, compared to the single-period setting, in a multi-period borrowing (B) regime the firm gains an additional ability to smooth investment (use the first-best input levels) by being able to borrow against future periods. However, the firm's borrowing is non-contingent ex-post (the same repayment must be made independent of cash flow) and hence, since we do not allow default, the firm faces an endogenous borrowing limit imposed by the non-negativity of consumption in the worst possible history of cash flow realizations. In the saving only (S) regime the firm can only use its savings to maintain the needed investment. In contrast, in the moral hazard (MH) regime the firm faces a single $t = 0$ budget constraint (the intermediary's zero profits constraint) in terms of expected inflows vs. outflows *both across time and across states of the world*. That is, through the financial intermediary, the firm can not only borrow against future cash flows (as it could do in the B regime) but also against other states of the world (unlike in the B and S regimes). In sum, in a multi-period setting, when the first-best is unattainable, the key difference between the B/S and MH regimes is enhanced in the direction already illustrated in our one-period example – namely, the firms' ability to carry resources across states of the world in the MH regime softens the debt repayment constraint.

To show the above formally, consider a two-period example. The saving only (S) and the saving and borrowing regime (B) feature six resource constraints indexed by time period and cash flow state history:

$$\begin{aligned} c_1^i + k_1 &= q_i + (1-\delta)k_0 + b^i - Rb_0 \;\; \text{for } i = \{L, H\} \\ c_2^{ij} + k_2^i &= q_j + (1-\delta)k_1 - Rb^i \;\; \text{for } ij = \{LL, LH, HL, HH\}, \end{aligned}$$

where the subscripts on $k$, $c$ and $b$ indicate time period and the superscripts indicate state. Resources can be moved across time via saving or borrowing (the $b^i$'s) but not across the states. In contrast, the moral hazard (MH) regime has a single ex-ante resource constraint – resources can be moved both across states and across time periods,

$$\sum_{i=L,H} \pi_1^i(k_1, z_1)(c_1^i + k_1 - q_i) + R^{-1} \sum_{ij=\{LL,LH,HL,HH\}} \pi_2^{ij}(k_2^i, z_2)(c_2^{ij} + k_2^i - q_j) = (1-\delta)(k_0 + R^{-1}k_1)$$

These findings have implications for the models' predictions about the joint distribution of the observables – capital $k$, investment $I$ and cash flow $q$, both in a cross-section as seen in Figure 1 for the one-period model, and over time and states as discussed above.

# 3 Data and methods

## 3.1 Data sources

We use two unique and rich data sources. Our first data source is the SABI-INFORMA database which collects annual public financial statements deposited by firms at the Spanish Mercantile Registry. Therefore, we have annual balance sheets and profit and loss statements for each firm. In addition, and crucial for our study, SABI-INFORMA provides static information on firm characteristics such as location, age, industry and, importantly, ownership.[28] The data allows us to obtain direct information about the firms' total assets ($A$), fixed assets ($k$), sales revenue ($S$) and age. Using the balance sheets, we also obtain information about firm investment ($I$), defined as fixed assets at time $t$ minus fixed assets at $t-1$ plus depreciation, that is, $I_t = k_t - (1-\delta)k_{t-1}$; firm cash flow ($q$), defined as operating profit plus depreciation; firm debt ($D$), defined as short-term creditors plus long-term creditors, liquidity ratio ($LR$), defined as cash / short-term creditors. We use the information about firm shareholders (ownership) to classify firms as family or non-family owned. Using ownership shares data, we are also able to identify whether each family firm is part of a larger network of firms.

Our second data source is the Credit Register (CIR) of *Banco de España* (the Spanish central bank). This database is unique because it is a census of the loans granted to Spanish firms by Spanish credit institutions (commercial banks, savings banks and credit cooperatives) or by the subsidiaries and branches of foreign banks operating in Spain.[29] The Bank of Spain Credit Registry (CIR) contains monthly information on the stock of all loans above a minimum threshold granted by credit institutions operating in Spain to Spanish firms. Since the minimum threshold was very low over our sample period (6,000 Euros), the available data covers essentially all bank loans granted to non-financial firms in Spain. CIR also provides information about the banking status of each firm (no relationship, single bank relationship, multiple bank relationships) and the amounts that each bank has lent to the firm.

Merging these two unique data sources provides us with annual information on ownership, banking relationships and economic and financial characteristics for a large sample of Spanish non-financial firms during the period 1997-2007. We refer to the merged data as SABI-CIR.[30] By construction this is an unbalanced panel, as

[28] Information is static in the sense that it reflects the status of the firm at the moment the information was first supplied, which may not always coincide with the actual firm status over all years.

[29] For detailed analysis of the CIR database see Jiménez et al (2006), (2009), (2012), (2014) and (2016).

[30] The list of all variables we use and their definitions is provided in Appendix B.

a consequence of entry and exit of firms, because of changes in the composition of the firms' portfolio for which SABI-INFORMA collects information, or due to our cleaning of the raw data.[31] We exclude publicly listed firms (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 the overall findings. The resulting dataset contains 4,749,653 observations, corresponding to 972,639 firms over the period 1997-2007.

A priori, the characteristics of our data (wide coverage in terms of firm age, size, region and industry) make it attractive for studying the presence and strength of financing constraints influencing firm investment patterns and cash flow sensitivity. Of particular interest to us are the roles of ownership and banking status in alleviating, or not, financial constraints on Spanish non-financial firms which, as non-listed companies, tend to be informationally opaque. We also use the firm ownership data to focus on family firms and networks of family firms (as formally defined in the next section) and understand the role played by family ties in softening financial constraints.

## 3.2 Firm Characteristics

### 3.2.1 Banking status

To investigate the role of banking relations in firms' investment decisions, we stratify the firms depending on their number of maintained banking relations. The number of banking relations for a given firm is defined as the number of banks reporting at least one loan to the Bank of Spain Credit Registry (CIR) for that firm and time period. According to this definition, in each year a firm in our sample can be classified as "unbanked", if it has no bank loans granted and registered in CIR, "single-banked" if it has loans from a single bank only and "multiple-banked", if it has loans by two or more banks.

### 3.2.2 Family and network status

To investigate the role of family firms, either directly or indirectly owned, we construct an indicator variable for each firm in our data, called 'family status', $f_j$ which reports whether firm $j$ is directly family-owned or part of family-based business network. The family status classification is done based on the date this information was requested from SABI-INFORMA (March 2014) and it is not possible to track any changes in this status over the sample period. We identified family firms using the following algorithm.

**Algorithm for defining family firms**

1. We start by creating a list of all non-financial firms in which 50% or larger share is held by an individual or family (reported shareholder type is "*una o más personas físicas o familias*"). For all firms in this list (280,534) we record their name and fiscal ID number. These firms are considered directly family owned.

2. To the list of directly family-owned firms constructed in Step 1 we add, using an iterative procedure, all other firms owned (with 50% or larger share) by firms in the family firm list. That is, we enlarge the initial set of family firms with all firms for which the major shareholders are family firms as defined in Step 1. To

[31] We dropped observations corresponding to firms declaring negative equity or interest payments higher or equal than total debt. We also dropped the value of a variable if the variable is by definition non-negative, but a firm reports negative value. We applied this filter to: sales, total assets, tangible assets, financial income, financial expenses, short term debt, long term debt, commercial debt, and cash. Only a small number of observations were affected and in a non-systematic way.

do so we use information on the holdings (names and percentage of shares) of non-financial firms in other firms.

3. We combine the 280,534 family firms from Step 1 with the ownership shares data from Step 2. We find 19,079 firms held (direct share $\geq$ 50%) by family firms from the Step 1 list. We check and drop any duplicate firms already included in that list (212 firms). As a result, at the end of this first iteration we add 18,867 firms to the list of family firms (it now contains 280,534 + 18,867 = 299,401 firms) held by individuals, families or by 9,839 family firms from the Step 1 list.

4. We then repeat the same process with the 18,867 firms added in Step 3, essentially proceeding along a tree and following the branches of ownership. The result is 3,497 additional family firms in the second iteration; 868 in the third iteration; 846 in the fourth; 8 in the fifth, 35 in the sixth and 0 in subsequent iterations. In the final list we drop duplicate observations (140) that correspond to firms held with equal percentage (exactly 50%) by two different firms. In these cases only one observation is recorded. The overall outcome from the recursive process is 304,515 family firms, either directly family-owned or indirectly owned by a family or individual via an ownership chain consisting of other firms.

In sum, we call *family firms* all firms that are controlled by an individual or family directly or through subsequent control stakes.

**Pure family firms vs. family-networked firms**

We further identify two sub-groups of interest within the set of family firms. Specifically, we call *pure family firms* all family firms which are directly owned by individuals or families and do not own other firms. In contrast, we call *family-networked firms* all family firms that are part of a network of firms constructed by ownership. The family-networked category consists of (i) all family firms that are directly owned by an individual/family and own other firms and (ii) all family firms that are owned by other family firms. The reason we distinguish these sub-groups of family firms is that they can look quite different in practice (e.g., a 'mom-and-pop store' vs. a firm which is part of a large business group).

In Figure 2 we show several examples of family-based networks among firms in our data. The simplest example of a network is *Case A*, where firm B is a subsidiary of firm A which is owned by family X. We are able to identify such direct relationships between firms as long as a firm such as A owns more than 50% of the shares of a subsidiary (firm B in the example). We do this up to the seventh level of ownership hierarchy (e.g., if firm B owns more than 50% of a third firm, and so on, up to a seventh firm).

More complex family networks in which a firm is owned by two or more family firms are also illustrated on Figure 2. We distinguish two possibilities: *Case B*, in which a single family firm (firm A) owns more than 50% of the shares of another firm (firm C); and *Case C*, in which two firms (A and B) jointly hold more than 50% of the shares of a third firm (firm C), but A and B individually hold less than 50% of firm C each. Case C is not considered in the current analysis given the identification complexity of such ownership structures.

To understand better the structure of family networks in our data, we present a case study in Figure 3. It was selected because of the multiple interactions it features, including both banked and unbanked firms. In Figure 3, an individual (Mr. X) is the sole owner of two firms (A, unbanked and B, banked), which are classified as pure family firms. Mr. X also holds 47% of a third firm (C, unbanked). If we only accounted for Mr. X's ownership, firm C would not be considered as family firm given our definition; however, when the ownership shares of other firms owned by Mr. X's family are taken into account, there is an additional 17% share, so the total adds up to 64% and hence firm C is classified as a family firm. Observe also that there are two more firms (D, unbanked and E, banked) which are not directly owned by Mr. X but are instead owned by firm C. This describes the second tier of the studied family network and consequently we classify D and E as family firms. All three firms C, D and E are family-networked family firms according to our definition. Those indirectly family owned firms may hold further shares in other firms. Indeed, this is the case for firm E which holds more than 50% share in another firm (F, banked). This relationship represents a third tier of the family network and firm F is also classified as family (-networked) firm. Finally, there are two more tiers in the studied case. Firm F is the sole holder of the unbanked firms G and H, while firm H owns 100% of firm K. Accordingly, firms G, H and K are also classified as family (-networked) firms, completing a five-tier family network.

In the Figure 3 case study there are further, more complex, ownership interactions involving firms I (banked) and J (unbanked). Both Mr. X and other family firms in the network own shares in I and J. However, only adjacent nodes with the criterion of $\geq$ 50% of ownership were considered in constructing the family networks from the whole data sample. Hence, we are not identifying cases such as firm J as a family-networked firm. In a sense, we are being conservative in our definition of family networks by counting only the cases in which we are completely certain that a given firm is fully controlled by a family. For example, firm C in case C on Figure 2 is also not included in the set of family firms because there is no guarantee that it is controlled more than 50% by family representatives. There are also data limitations regarding such cases because identification numbers or exact shares are missing for minor shareholders.

Notice that, within the same family network, unbanked and banked firms coexist. As mentioned earlier, this can help us understand financing connections within the network. For example, unbanked family firms may benefit from funding from other firms in the family network which are either single or multi-banked. In the case study depicted on Figure 3, unbanked firms such as G, H, J and K may benefit from direct connections with banked firms such as B, E, and F. Banked firms could also receive additional financing through the network. Such family network interactions may therefore explain our finding that family firms seem to be facing less strict financing constraints compared to non-family firms. But we will find that bank relationships per se do not always offer a strict monotonic improvement and so being indirectly connected to banks should not either.

To further clarify the different categories of firms defined by direct or indirect family ownership, consider the 2004-2007 balanced panel sample. Out of 100,745 firms in total, 42,012 (41.7%) are family firms. Within the set of family firms, 39,620 (94.3%) are directly owned by a person or family and 39,715 (94.5%) do not own other firms. Taking the intersection of these two sets yields 38,950 (92.7%) pure family firms. Also within the set of family firms, 2,297 firms (5.5%) are owners of other firms, out of which 670 are directly owned, via 50% or higher share by a person or family. The latter, together with the 2,392 firms (5.7%) owned by family firms, comprise the total of 3,062 (7.3%) family-networked firms. The relatively large fractions of firms owning or owned by other firms within the family-networked category suggest that many family networks are relatively complex.

### 3.3 Sample construction and summary statistics

To structurally estimate and compare the three models of dynamic financial constraints from Section 2, we use data on firm's tangible fixed assets, $k_{jt}$ which we interpret as firm size; investment $I_{jt}$; and net profits, $q_{it}$, where

$j$ indexes firms and $t$ indexes time. We therefore keep only the firms $j$ for which data on all three variables $(k_{jt}, I_{jt}, q_{jt})$ are available for at least one year. This reduces the sample from 4,749,653 observations on 972,639 firms to 3,718,286 observations on 805,118 firms, a drop of approximately 20%. The reason is data availability in the non-mandatory (unlike CIR) SABI-INFORMA database – we need information on all three variables and moreover computing firm investment requires two observations on fixed assets from consecutive years.

Table 1, part A, column (1) reports summary statistics for all firms with non-missing observations of $k, I$ and $q$. The mean firm age is 11 years. A large fraction (48%) have loans from multiple banks, 31% are single-banked and a significant fraction (21%) are unbanked. Based on our algorithm, 32% of the firms are classified as family firms, of which the majority are pure family firms. The summary statistics for firm assets, investment and cash flow reveal substantial heterogeneity in all three variables – the standard deviations are much larger than the mean values and there are long right tails. For example, the median investment is approximately 10,000 Euros, which is nearly 20 times lower than the mean investment (196,000) and nearly 1,500 times lower than the standard deviation of investment across firms.

Consistent with the model assumption that firms are in steady state with respect to their financial environment, in the structural estimation we restrict attention only to firms which maintain a *continuing banking status*, that is, continuously unbanked, continuously single-banked, or continuously multi-banked over all years in which they are present in the full dataset. This reduces the sample to 1,493,181 observations on 385,539 continuing banking status firms – see column (2) in Table 1, part A. Comparing columns (1) and (2), continuing banking status firms are on average larger (with higher mean and median fixed assets $k$, investment $I$ and cash flow $q$), however, because of the large standard deviations, the differences in the means are not statistically significant. Mean firm age and the share of family firms are approximately the same in columns (1) and (2) but there is a larger fraction of multi-banked and a smaller fraction of single-banked firms among continuing banking status firms. We do robustness runs with firms that switch banking status (see Section 4.3).

In Section 4, in order to fully test the implications of our dynamic model we perform maximum likelihood estimation runs with panel data on investment and cashflow. To do so we construct three balanced panels of capital, investment and cash flow, $\{k_{jt}, I_{jt}, q_{jt}\}$ data: a panel covering 1997-2000 (the first four years in the data), a panel spanning the intermediate period, 2001-2003, and a panel covering 2004-2007 (the latest years for which we have data). We also do runs with cross-sectional data. Each of the three balanced panels consists of all continuing banking status firms with non-missing data on $\{k_{jt}, I_{jt}, q_{jt}\}$ for all the panel years. Using the $q$ and $k$ data we estimate the production function $P(q|k, z)$ in each of the three panels separately (see Appendix A2 for details). We find that the production relationship remains very stable over the time periods.

The main reason for breaking the data into sub-periods and our choice of the panels' length is the trade-off between having a sufficiently large sample size and having a balanced sample, with multiple observations of the same firm over time. We need a sufficiently large sample size to be able to stratify by firm size, age, region, industry, banking, and family status. A longer panel reduces the sample size and can introduce selection from attrition. We need a balanced sample of firms to be able to do estimation runs with multi-period panel data, that capture the joint distribution of capital, cash flow and investment both across firms and over time. A second reason for breaking up the data this way is to allow us to capture possible changes in the financial conditions of the various categories of firms over the different sub-periods we study, as described earlier. We do perform robustness estimation runs with continuing banking status firms but without requiring a balanced panel (the firms in column 2), and show that our results remain robust (see Section 4.3).

The three balanced panels contain 880,727 observations for 135,451 distinct firms in total. These firms represent 13.9% of all firms in the SABI-CIR database and account for approximately 43% of aggregate total assets, 46% of aggregate debt, and 55% of aggregate bank debt in SABI-CIR. Compared to the entire population of Spanish non-financial firms, the coverage of the balanced panels is around 44% of aggregate total assets and 59% of the total loan amount given by Spanish credit institutions.

Columns (1) and (2) in Table 1, part B report summary statistics for only the latest period 2004-07, as an example, with corresponding statistics for the 2004-07 balanced panel of continuing banking status firms displayed in column (3).[32] Comparing column (3) with column (1): all firms and column (2): all continuing banking status firms, we see that the firms in the 2004-07 balanced panel (column 3) are larger on average and there are larger fractions of multi-banked and family firms and smaller fractions of unbanked and single-banked firms.[33][34] However, despite these differences, our robustness runs with respect to the sample construction in Section 4.3 show that our main result that family firms are less financially constrained than non-family firms remains robust.

Tables 2A-2E contain summary statistics and financial indicators for the firms in the three balanced panels, pooled and stratified by different observable characteristics. In Table 2A we stratify the firms by ownership and banking status. Comparing family and non-family firms, the largest fraction of family firms is among multi-banked firms (around 50% of all), while the smallest fraction is among unbanked firms (17%). Comparing median assets, cash flow and investment, family firms tend to be larger than non-family firms, although the difference is small for multi-banked firms. In addition, the financial indicators section of Table 2A (not used in the structural estimation) suggests that family firms have larger median sales and cashflow-to-assets ratio compared to non-family firms.[35]

Table 2B breaks down family firms into pure family vs. family-networked, as defined above, and compares these categories to non-family firms. Pure family firms appear more similar to non-family firms in most reported dimensions. Family-networked firms tend to have larger median and mean assets, cash flow and investment. There is larger variability (comparing the coefficient of variation) in assets and investment among non-family firms compared to pure family and family-networked firms. Table 2B also reveals features of the financial situation of family-networked firms that are consistent with our estimation results in Section 4. Relative to pure family and non-family firms, family-networked firms have lower median liquidity ratios, especially among the unbanked, i.e., they seem to need less liquidity. Family-networked firms also have consistently lower debt-to-assets ratios, that is, they appear less reliant on formal funding.

---

[32] Corresponding summaries for the 1997-00 and 2001-03 periods are provided in Table A1 in the Appendix. The statistics in Table A1, column (2) for the 2001-03 period are very similar to those in 2004-07, while in 1997-00 we observe slightly larger median values of $k, I, q$ which could be explained by the larger fraction of multi-banked firms (72% vs. 51% in Table 1).

[33] Because of the large standard deviations, the differences in mean capital, cash flow and investment are not statistically significantly different from those in columns (1) and (2); the differences in firm type fractions are statistically significant.

[34] The mean firm age in column (3) is also larger, but this is in part by construction, since the reported mean age is computed at the observation (not the firm) level and we require each firm in the balanced panel to have observations for at least four consecutive years.

[35] Across banking status, unbanked firms tend to be the smallest in terms of mean and median assets $(k)$, cash flow $(q)$ and investment $(I)$ but the differences in the means are not statistically significant because of the large standard deviations. Comparing the standard deviations we see more heterogeneity among multi-banked firms. Multi-banked firms also have larger mean and median total assets and sales, higher debt-to-assets ratio, lower liquidity ratio and are older compared to unbanked and single-banked firms.

### 3.4 Regional and sector heterogeneity

Firm size, ownership structure and the number of banking relationships involve business decisions that can respond to other basic observable characteristics such as sector (industry) and location. To account for such possible endogeneity we also stratify the data by major industry and region. We then estimate the financial regimes in each of these stratifications separately (see Tables 5C and 5D). In terms of industry, we divide firms into “construction” (those with primary business activities in construction or real estate) and “non-construction” (all other non-financial firms). In terms of location, we divide firms by defining two broad geographic regions in Spain. Region 1 corresponds to the high-growth parts of Spain, the Mediterranean, Madrid and the Ebro river, while Region 2 corresponds to the rest of the country, more agricultural, less developed and aging. It is practically infeasible to stratify the data into finer divisions by location and industry for computational time reasons (each set of runs takes up to a month) and sample size (when we further stratify by banking and family status), however, based on the extensive policy and research experience of one of us at the Bank of Spain, these broad categories represent well the natural dividing lines in the Spanish economy in the period we analyze.

In Table 2C we compare the fractions of firms and the mean, median and standard deviations of the data variables used in the estimation for construction vs. non-construction firms. We find evidence of heterogeneity by industry. There are relatively more construction firms among the unbanked (31%) vs. among the multi-banked (20%). There is also a higher fraction of family firms among non-construction firms. Unbanked and single-banked construction firms have larger mean capital stock and investment than non-construction firms but the opposite is true for multi-banked firms. However, the differences in the means are not statistically significant because of the large standard deviations. The coefficient of variation of cash flow is much larger for non-construction unbanked and multi-banked firms than for construction firms of the same type and the coefficient of variation of investment is larger for non-construction multi-banked firms. However, these statistics also pick up cross sectional variation over a mix of non-construction industries.

In Table 2D we compare firm characteristics by region, as defined above. As in the stratification by industry, we find evidence for regional heterogeneity in our data. Region 1 firms have larger average capital stock, cash flow and investment and also higher standard deviations of these variables compared to Region 2 firms. The differences in the means are not statistically significant because of the large standard deviations. There are relatively more family firms among Region 2 firms.

To address possible concerns about differential access to credit (market structure) across locations we computed a Herfindahl concentration index (HHI) for the two regions we define and performed a statistical test on whether the difference in bank concentration is significant.[36] The results show that we cannot reject the null hypothesis that both regional HHIs are equal for 1998 and 2007, the beginning and end of the time period we analyze (p-values 0.42 and 0.23 respectively). Indeed, in that period there was a lot of bank competition throughout Spain.

Finally, in Table 2E we break down and summarize the distribution of firm observations in our data by banking status, size and age, family status, region and industry. There are small differences in the regional composition among firms with the same banking relationship status. We also see heterogeneity by industry, with relatively more unbanked and less multi-banked among construction firms compared to among non-construction firms. These differences in the fractions of firms are statistically significant given the large sample sizes.

[36] We thank an anonymous referee for this useful suggestion.

### 3.5 Computation and empirical method

To solve and structurally estimate the three dynamic models of firm investment under alternative financial constraints, we re-write each of the dynamic programming problems from the previous section as linear programs (Prescott and Townsend, 1984; Phelan and Townsend, 1991; Karaivanov and Townsend, 2014). We first discretize the problems in Section 2.2 by assuming that all state and choice variables take a finite number of values. We then use the discretized problems to define the joint probability distribution of the choice variables ($z, k'$ and $c$) and cash flow $q$, conditional on the current state, as new choice variables in a linear program. Details are available in Appendix A.

Our rationale for using linear programming (LP) is two-pronged. First, it is well-known from the literature (e.g., Rogerson, 1985; Phelan and Townsend, 1991) that the incentive compatibility constraints in moral hazard problems can cause non-convexities. One possible way to deal with this issue is to make specific assumptions on the preferences, the production function or the distribution of shocks and use the "first-order approach" replacing the incentive constraint by its first order condition (Rogerson, 1985). Alternatively, Abraham and Pavoni (2008) propose a numerical algorithm whereby a candidate solution is obtained using the first-order approach and then a global maximization routine is used to verify whether an optimum is indeed achieved.

In contrast, by writing the contracting problems as linear programs in terms of the joint probabilities (lotteries) over choice and state variables, we obtain problems that are convex by construction and so we can use arbitrary (including non-convex or non-separable) preferences or technology. While the saving only and borrowing regimes can be solved by non-linear methods, we use the LP method to ensure consistency with the solution to the moral hazard problem, which is important for the model comparison tests we perform. A second important advantage of our numerical approach is that it offers a direct and intuitive mapping from the LP solutions, $\pi$ (joint probability distributions) to the *likelihood* of the observed data – capital $k$, investment $I$, and net profits $q$, under the null hypothesis of each financial regime.

To structurally estimate and statistically test across the three models: saving only (S), non-contingent saving/borrowing (B) and moral hazard (MH), we use the approach developed in Karaivanov and Townsend (2014), adapted to the specific case here. In short, we use the probabilistic form of the solutions to the dynamic linear programs exhibited in Appendix B, together with assumed parametric distributions for measurement error and the initial unobserved state ($b$ or $w$) and construct a likelihood function between the data and each of the three alternative dynamic models of firm finance. A basic description of our empirical method is provided in Appendix C and interested readers are referred to Karaivanov and Townsend (2014) for the full details.

We then use the MLE parameter estimates and likelihood value for each regime in a model comparison test (Vuong, 1989). We approach the data as if we are agnostic about which model fits best and let the data themselves determine this. The Vuong test shows whether we can statistically distinguish between any two models based on their fit with the data, as in a model 'horse race'. The test does not require that either of the compared models be correctly specified. In our setting the compared models are statistically non-nested (see Vuong, 1989 for formal definition) and the test statistic is normally distributed under the null hypothesis that the two compared models are equally close to the data in the KLIC sense. If the null is rejected (the Vuong test statistic is sufficiently large), we conclude that the model with larger likelihood value is closer to the data than the alternative.[37]

[37] Due to the Vuong test's pairwise nature, its outcomes are not always fully transitive (for example, models X and Y can be each tied with model Z but X can win over Y). When reporting results, we handle this issue by listing as 'best fitting' the highest-likelihood model

# 4 Results

## 4.1 Overview of main results

We give an overview of our baseline estimation results in Table 3 in which we report the frequency of the best-fitting estimated financial regimes across all data stratifications we consider – by firm age and size, banking status, family ownership status, region and industry. When computing these frequencies we counted as $1$ each time a regime is best-fitting and not tied with others and as $1/n$ each time a regime is tied in an $n$-way tie ($n = 2$ or $3$). The complete estimation results are reported in Tables 4, 5A-5D and 6 below and we provide detailed discussion related to each table in Section 4.2.

When interpreting the MLE Vuong test results we keep in mind the theoretical prediction that the moral hazard (MH) regime is the least financially constraining of the three settings we study, as it enables firms to transfer resources both across states of the world and time. The borrowing and saving (B) regime is more restrictive, since indebted firms must make non-contingent payments. Finally, the saving only (S) regime is the most financially constraining as firms can only use buffer stock savings to react to cash flow fluctuations.

**Table 3 - Vuong test winning regime frequency summary**

| | family firms | | | non-family firms | | | source |
|---|---|---|---|---|---|---|---|
| best-fit regime, % of all runs | S | B | MH | S | B | MH | table |
| all | 17% | 63% | 20% | 65% | 24% | 11% | 5A |
| small and young | 41% | 54% | 4% | 67% | 33% | 0% | 5B |
| large and old | 0% | 67% | 33% | 46% | 39% | 16% | 5B |
| unbanked | 39% | 56% | 6% | 78% | 22% | 0% | 5A |
| single-banked | 0% | 78% | 22% | 83% | 17% | 0% | 5A |
| multi-banked | 11% | 56% | 33% | 33% | 33% | 33% | 5A |
| construction | 13% | 87% | 0% | 63% | 37% | 0% | 5C |
| non construction | 13% | 67% | 20% | 63% | 37% | 0% | 5C |
| region 1 | 10% | 77% | 13% | 43% | 57% | 0% | 5D |
| region 2 | 7% | 87% | 7% | 63% | 30% | 7% | 5D |
| pure family | 35% | 52% | 13% | n.a. | n.a. | n.a. | 6 |
| family networked | 17% | 49% | 34% | n.a. | n.a. | n.a. | 6 |

Table 3 reveals three major patterns on which we elaborate in Section 4.2.

1. family firms are estimated to be less financially constrained than non-family firms, with more frequency/weight on the MH and B regimes as best-fitting and less frequency on the most constrained saving only regime S.

and all other models statistically tied with it, in order of decreasing likelihood.

This holds in the whole sample and in each stratification by age/size, banking relationship, industry and region. In most stratifications of non-family firms the moral hazard regime is always dominated by the saving only and borrowing regimes.

2. consistent with the literature, we find a clear pattern by firm age and size – the best-fitting financial regime for large and old firms is less constrained than the best-fitting regime for small and young firms (larger frequency of MH and smaller frequency of S as best fitting among large and old firms).

3. within the category of family firms, pure family firms are estimated to be more financially constrained (with higher frequency on the more constrained regimes S and B) than family-networked firms – the most flexible MH regime is best-fitting in 34% of all runs among family-networked firms vs. only 13% of all runs among pure family firms, the opposite is true for the most restrictive saving only regime.

In addition, the differences in the best-fitting regime percentages in Table 3 are not statistically significantly different across the two regions and the two industry stratifications. Comparing the Vuong test results across the firms with different number of banking relations, we also see higher frequency of the most constrained regime (S) among unbanked firms (and also among single-banked non-family firms) and higher frequency of observing the MH regime as best-fitting among multi-banked firms.[38]

**Computed example**

To interpret quantitatively the significance of a firm being in a different financial regime, we compute an example, using a representative set of parameter estimates (Table 5A, row 3.6), to solve the three regimes and show the present-value gain in expected firm payoff (profits) for a firm in the saving only (S) regime from switching to the borrowing (B) regime or switching to the moral hazard constrained regime (MH). The gains are computed for each capital level $k \in K$ where $K$ is the five-point capital grid used in the linear programs (see Appendix A2 for details). Figure 4 shows that the gains can be significant (up to 20% for MH and 10% for B) at low capital stock levels, at which the firms are most constrained.[39] In the data from which the parameters used in Figure 4 were estimated, 55% of all firms have capital stock that maps to the three lowest $k$ levels (with MH gains relative to S of over 15%) and additional 27% of firms have capital stock that maps to the fourth largest level ($k = .235$).

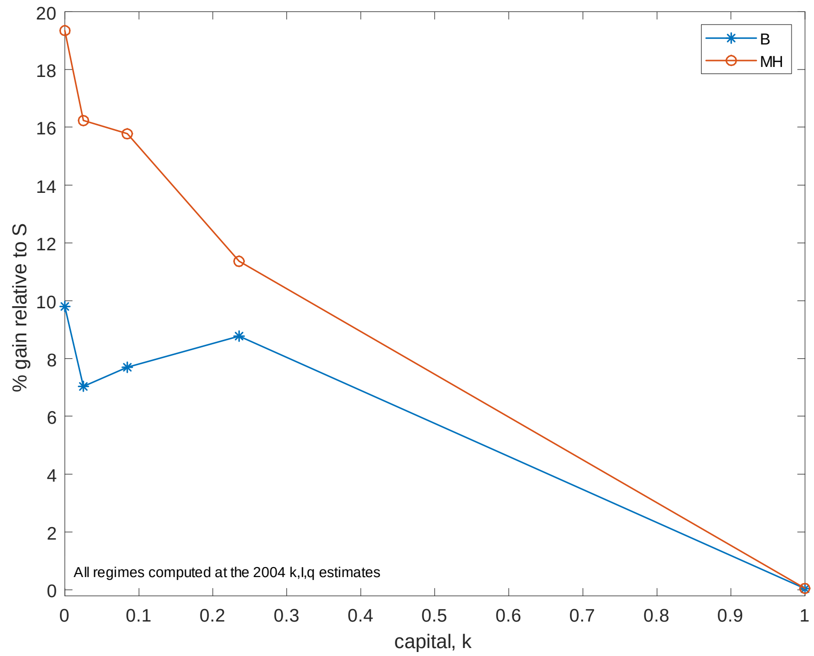
## 4.2 Model comparisons

Tables 4–6 report the full details of our baseline Vuong test model comparison results. We display the results in reverse chronological order, starting from the most recent data. Each table column contains the Vuong test statistics from comparing the three financial regimes pairwise. A statistically significant test statistic (we report 1%, 5%, and 10% significance levels, denoted by $^{***}$, $^{**}$ and $^{*}$ respectively) indicates that we can reject the null hypothesis that the compared models are equally close to the data in KLIC sense. For each table entry, the model regime with abbreviation (MH, B or S) reported in the table fits the data better than the alternative model to which it is compared. The last column of each table section reports the best-fitting model overall for that particular estimation run. All results assume risk neutrality ($\sigma = 0$). We perform a robustness test allowing for risk aversion in Section 4.3.

[38] The reported frequencies are based on the results in Table 5A.

[39] We also did several other runs, with different MLE estimates, and found gains from switching to the B regime as low as 7% for low $k$ and gains from switching to the MH regime as high as 32%.

Figure 4: Example: present-value gains relative to saving only (S)




#### 4.2.1 Age, size and banking status

In Table 4, we study how the best-fitting financial regime varies by firm age and size by comparing all firms vs. small and young firms vs. large and old firms. For each balanced panel (e.g., 2004-07), we define "small and young" firms to be all firms with fixed assets $k$ below the median and which are less than 10 years old in the first panel year (e.g., 2004). Similarly, we call "large and old" all firms with assets $k$ equal or above the median and which are 10 years or older in the first panel year. We estimate with both panel data (on investment, rows 1.1-1.3; or investment and cash flow, rows 2.1-2.6) and with cross-sectional data on assets, investment, cash flow (parts 3-5 in Table 4).

The model comparison results in Table 4 reveal a clear pattern by firm age and size – the estimated best-fitting financial regime for large and old firms is less constrained than the best-fitting regime for small and young firms. Across all Table 4 specifications, we count 16 instances of strictly less constrained best-fitting regime for the large and old firms (MH vs. B or S, or B vs. S), 1 instance of weakly less constrained regime (row 5.5), and 6 instances in which the estimated best-fitting regime is the same. In only one case (row 5.6) the result is mixed – regime B fits best for large and old firms while MH,B and S are tied for small and young firms. These results, suggesting a less constrained financial environment for large and old firms are consistent with the previous findings in the empirical literature that older and larger firms can achieve less constrained investment / cash flow relationship compared to small and young firms. In our data this is especially evident in the 2004-2007 pre-crisis period when the lending boom and activity was at its peak.[40]

[40]The MH regime is best-fitting (including one tie) among large and old firms in 11 out of the 15 estimated specifications with 2004-07 data. Using $(I, q)$ panel data (rows 2.1-2.6) all six specifications yield MH as best-fitting among large and old firms while S (and B in two

Regarding the role of banking relationships, our results depend on the firm size and age. In the whole sample (Table 4, columns A), the best-fitting financial regime for unbanked firms tends to be (weakly) more constrained than the best-fitting regime for single- and multi-banked firms – for example, S for unbanked vs. B for single-banked and multi-banked firms in rows 3.1–3.3 and 3.4–3.6; or S for unbanked, B for single-banked and MH for multi-banked firms in rows 2.1-2.3 and 2.4-2.6. The results for small and young firms (Table 4, columns B) are similar – the data on unbanked firms are best fit by a (weakly) more constrained financial regime (S) than the corresponding data on single- or multi-banked firms. This is evidence that small and young unbanked firms are more financially constrained. In contrast, among large and old firms (Table 4, columns C) we often see a different pattern – the best-fitting financial regime for unbanked firms is the same or less restrictive (for example, MH in rows 3.1, 3.4, 4.1) than that for the corresponding single- and multi-banked firms.

We summarize the results about the best-fitting regime by firm age and size and banking status in Table 4B. The reported frequencies are computed the same way as in the summary Table 3 in Section 4.1. Large and old unbanked firms have the highest frequency of the least constrained financial regime (MH) as best fitting, 63%.[41] In contrast, in the whole sample and among small and young firms, unbanked firms are estimated to be more constrained, with respectively 75% or 88% of all estimation specifications yielding the saving only, S regime as best fitting.

**Table 4B - Summary of the best-fitting regime in Table 4**

| | best-fitting regime frequency | | |
|---|---|---|---|
| | S | B | MH |
| whole sample, unbanked | 75% | 25% | 0% |
| whole sample, single-banked | 25% | 75% | 0% |
| whole sample, multi-banked | 0% | 63% | 38% |
| large and old, unbanked | 13% | 25% | 63% |
| large and old, single-banked | 13% | 44% | 44% |
| large and old, multi-banked | 0% | 63% | 38% |
| small and young, unbanked | 88% | 13% | 0% |
| small and young, single-banked | 81% | 19% | 0% |
| small and young, multi-banked | 48% | 48% | 4% |

### 4.2.2 Family vs. non-family firms

In Tables 5A-5D we stratify the data by family status, that is, whether a firm is owned or not by a person or family, either directly or through a network of subsidiaries, in addition to stratifying by firm size, age and banking status.

cases) are best-fitting for small and young firms.

[41] A possible explanation is these firms' different composition by industry (53% of the large and old unbanked firms are in the construction and real estate sector, compared to 29% for large and old single-banked, 17% for large and old multi-banked and 31% for all unbanked).

In Table 5A, we see that the best-fitting financial regime for family firms is less constrained (allowing for the tie in row 1.3) compared to the best fitting regime for non-family firms. Comparing across columns A and B row by row, there are 13 instances of strictly less constrained financial regime (MH vs. B or S or B vs. S) for family firms, 3 instances of weakly less constrained regime due to ties (rows 2.4, 3.2, 3.6) and 10 instances with the same best-fitting regime. This pattern holds both with investment and cash flow or investment panel data (sections 1, 2 and 4 in Table 5A), and also with cross-sectional $(k, I, q)$ data from each of the three balanced panels. The most restrictive saving only regime appears as best fitting 18 times for non-family firms vs. only 5 times for family firms.

The same pattern, of family firms being less financially constrained, holds when we control for firm size and age in Table 5B (compare the sections labeled "Small and young" and "Large and old"). There are 10 instances of strictly less constrained regime (MH vs. B or S or B vs. S) for family firms (7 of which among large and old firms); 6 instances of weakly less constrained regime (e.g., B vs. B,S) and 13 instances of the same best-fitting regime. [42]

In Table 5C we stratify the firms in our sample into "construction" vs. "non-construction", as defined in Section 3, to see if our results are sensitive to possible industry differences. We see in Table 5C that our main result, that family firms are less financially constrained than non-family firms, holds in each industry sub-sample. Specifically, there are 13 instances of strictly less constrained (MH vs. B or S or B vs. S) best-fitting financial regime for family firms (5 of which among construction firms), 8 instances of a weakly less constrained regime (ties), and 9 instances of the same best-fitting regime.

In Table 5D we stratify the firm data into two major geographic and economic regions, as defined in Section 3. Once again we see that, in both the faster-growing and more developed Region 1 and the more agricultural and less developed Region 2, the Vuong test results indicate that family firms are less financially constrained than non-family firms. There are 13 instances of strictly less constrained best-fitting financial regime for family firms (MH vs. B or S, or B vs. S), 3 instances of a weakly less constrained regime and 13 instances of the same best-fitting regime.[43]

Overall, across all our baseline estimation runs (the 117 specifications in Tables 5A–5D) there is no exception from the pattern that, allowing for Vuong test ties, (one of) the best-fitting financial regime(s) for family firms (columns A) is weakly less constrained than the best-fitting regime for non-family firms with the same banking status, age, size, industry or region in the same year (columns B).

### 4.2.3 Pure family vs. family-networked firms

To further investigate the mechanism behind our results with family firms, in Table 6 we divide family firms into two categories, as defined in Section 3.2.2: "pure family" firms, directly owned by individuals or families and "family-networked" firms, belonging to a network of family firms constructed by ownership. Columns A in Table 6 report the Vuong test model comparison results for pure family firms, columns B report the Vuong test results for family-networked firms and columns C report the results for non-family firms (repeating the results from Table

[42] One specification with 1997 cross-sectional data (row 9.6) is the only partial exception, with a three-way tie (MH,B,S) for best fit for family firms and regime B best-fitting for non-family firms.

[43] Region 1, 1997 single-banked firms is the only partial exception in which a more constrained regime (S) is tied for best-fit for family firms.

5A, to make comparisons easier).

There are two main takeaways from Table 6. First, comparing the corresponding entries in columns A and B, we find that family-networked firms are estimated to be less financially constrained than pure family firms. There are 8 instances of statistically significantly strictly less constrained financial regime for family firms (e.g., MH vs. B or S; B vs. S), 4 instances of weakly less constrained regime (e.g., B,S vs. S) and 3 instances of the same best-fitting financial regime. There are seven cases with a three-way tie (MH,B,S) in which a more constrained regime (S) is statistically tied for best-fitting for family-networked firms (e.g., row 1.3 or 2.3) but in all these cases a weakly less constrained regime (MH) is also best-fitting and achieves the highest likelihood.

The second takeaway from Table 6 is that pure family firms tend to fall in between the non-family and family-networked firms and in many specifications seem more similar to non-family firms than to family-networked firms.[44] Comparing columns A and C in Table 6, there are 13 specifications in which the best-fitting regime is the same for pure family and non-family firms. In the remaining cases (with one exception in row 1.3) pure family firms exhibit a less constrained financial regime (e.g., MH vs. S or B vs. S) than non-family firms.

#### 4.2.4 Discussion

Unfortunately, because of data limitations we cannot complement or re-do Tables 5B-5D with comparisons between family-networked firms vs. all other firms stratified by size and age, industry, region and banking status. The reason is sample size – when broken down by banking status and/or size and age, etc., the sub-samples of family-networked firms become very small (as low as 200 firms) which makes the Vuong tests inconclusive. We see evidence of this in Table 6, columns B (observe the large number of regime ties, including three-way ties, relative to columns A and C). We were able, however, to compare pure family vs. family-networked vs. non-family firms stratified by age and size alone, for any continuing banking status (rows 6.1 and 6.2 in Table 6). The results match our findings from the rest of Table 6 – family-networked firms are (weakly) less financially constrained than non-family firms while pure family firms look the same as non-family firms. These findings also re-affirm our previous results from Table 4, that large and old firms tend to be less constrained than small and young firms.

Regarding the underlying mechanism through which family-networked firms seem able to overcome financial constraints better than pure family and non-family firms, a potentially important caveat is that, given the available data, we are unable to directly study possible peer effects or spillovers within the family networks. In our data, we look at each individual firm and determine its banking status, age and size. In a network, however, family firms are linked and another firm in the network could have better financial access that benefits all members. However, we do account for firm-level characteristics on which we have available data (size, age, industry, region, banking status) and our results remain robust.

### 4.3 Robustness

In Table 7 we perform additional estimation runs to check the robustness of our main results. We first investigate the robustness of our results with respect to the sample of firms used in the estimation and model comparisons. In Table 7, rows 1.1-1.3, instead of using only firms belonging to 2004-07 balanced panel (as in Table 5A), we use *all*

[44] As an example, see row 2.2 using $(I, q)$ panel data, where the B regime fits best for pure family firms, the S regime for non-family and the MH regime for family-networked firm.

*firms* with available 2004 $(k, I, q)$ cross-sectional data.[45] The Vuong test results show that our main conclusions remain robust – the data from family firms are fit best by a less constrained financial regime (MH vs. B or S or B vs. S) than the data from non-family firms. Similar results obtain when using 2000 $(k, I, q)$ cross-sectional data (Table 7, rows 1.4-1.6).

In Table 7, rows 2.1-2.2 we use 2004 and 2007 $(k, I, q)$ cross-sectional data and estimate the model on firms that switch their banking status over time at least once. In these robustness runs we do not require continuing banking status as in our baseline runs. Our main result that family firms are less constrained than non-family firms still holds albeit in a weaker form. Note, however, that this robustness run stretches the interpretation and connection between the estimation sample and the theory, as we do not allow transitions in banking status or financial regime when we compute the financial regimes.

Next, (Table 7, rows 3.1-3.3) we pool all firms with any continuing banking status in the 2004-07 panel (that is, we pool the unbanked, single- and multi-banked firms together) and show that our main results from Table 4 (on firm size and age) and Table 5A (on family status) remain valid. Large and old firms are less financially constrained (MH or B are best-fitting) than small and young firms and family firms are not more constrained than non-family firms.

In Table 7, rows 4.1-4.3 we perform a "placebo test" by assigning family status *at random* to the firms in our sample, using 2004 $(k, I, q)$ cross-sectional data. As expected, the best-fit financial regimes are identical across columns A (pseudo family firms) and columns B (pseudo non-family firms). Moreover (not shown in the Table), the maximized likelihoods for each model are numerically very similar between the two placebo samples.

The next set of robustness checks address the model assumptions about risk preferences and technology. In Table 7, rows 5.1-5.3 we allow for a small degree of risk aversion by setting the relative risk aversion parameter to $\sigma = .1$. Our main result regarding family vs. non family firms is preserved – the best-fitting financial regime for family firms is weakly less constrained than the best-fitting regime for non-family firms.

In rows 6.1-6.3 we introduce quadratic capital stock adjustment costs in the model of the form $g(k_t, I_t) = \frac{1}{2} b k_t \left[ \frac{I_t}{k_t} \right]^2$ and estimate the added parameter $b$. Adding adjustment costs blurs the statistical distinction between the financial regimes both across banking status and family status. The reason is that, unlike most of the literature, we are not comparing a complete markets model vs. a model with adjustment costs as stand-in for financial or other frictions. Instead, we model the source of financial frictions explicitly, as either arising from exogenously incomplete markets (the S and B regimes) or endogenously incomplete markets (the MH regime). Adding adjustment costs in our three financial regimes thus dilutes, in an exogenous way, the distinction across the financial constraints we model.

Finally, In Table 7, rows 7.1-7.3 we do a robustness check on the model timing by using the joint distribution of *next-period* assets, investment and cash flow, that is $(k', I, q)$ instead of $(k, I, q)$. The data on family firms is again best-fit by a less constrained financial regime than the data on non-family firms, confirming our earlier findings.

[45] This yields a significantly larger sample size (for instance, there are 6,337 unbanked family firms in Table 7, row 1.1 vs. 3,221 unbanked family firms in the 2004-07 balanced panel used in Table 5A, row 3.1).

## 5 Conclusions

We estimate using maximum likelihood three prototypical non-nested dynamic models of financial constraints ranging from saving only, to exogenously incomplete credit market (borrowing/lending in a single asset), to endogenously incomplete market with state-contingent borrowing and transfers (moral hazard) for firms differing in their number of bank relations, family ownership status, age, size, region and industry. In line with our dynamic models of firm investment subject to a persistent source of financial constraints, we focus attention to firms that maintain the same banking and family ownership status continuously. We then determine the type of financial constraints for each firm category using a statistical model comparison test. The unique combination of our two data sources allows us to disentangle the roles that firm characteristics such as bank/non-bank relationships and family/non-family ownership play together with firm age, size, region or industry and to ultimately obtain a score card for which firm categories are more severely financially constrained and which less so.

We find strong evidence that family firms, directly or indirectly owned, are less financially constrained than non-family firms (the estimated best-fitting regime is MH for family vs. B or S for non-family or B vs. S respectively). Moreover, we find clear evidence that firms which are part of family-based networks constructed using ownership shares face less strict credit constraints compared to non-networked ('pure family') firms that are family-owned but do not own other firms. These results are robust across firm size and age stratifications and across banking status, region and industry. We interpret these findings as showing that family-based firm networks appear to reduce or surmount financial constraints that otherwise may limit firm investment, growth and ability to smooth out cash flow shocks.

## References


[1] Abraham, A. and N. Pavoni (2008), "Efficient Allocations with Moral Hazard and Hidden Borrowing and Lending: A Recursive Formulation," *Review of Economic Dynamics* 11(4), p.781-803.

[2] Ahlin, C. and R. Townsend (2007), "Using repayment data to test across models of joint liability lending", *Economic Journal* 117, p.F11-F51.

[3] Albuquerque, R. and H. Hopenhayn (2004), "Optimal Lending Contracts and Firm Dynamics", *Review of Economic Studies*, 71(2), p.285-315.

[4] Ales, L. and P. Maziero (2016), "Non-exclusive Dynamic Contracts, Competition and the Limits of Insurance", *Journal of Economic Theory* 166, p.362-95.

[5] Alessandrini, P., F. Presbitero and A. Zazzaro (2009), "Banks, Distances and Firms' Financing Constraints", *Review of Finance* 13, p.261-307.

[6] Allen, F. (1985), "Repeated Principal-Agent Relationships with Lending and Borrowing", *Economic Letters* 17, p.17-31.

[7] Alonso-Borrego, C. (1994), "Estimating dynamic investment models with financial constraints", Working paper 9418, CEMFI.

[8] Anderson, R., S. Mansi and D. Reeb (2003), "Founding family ownership and the agency cost of debt", *Journal of Financial Economics* 68(2), p.263–85.

[9] Andrés, C., (2011), "Family ownership, financing constraints and investment decisions", *Applied Financial Economics* 21, p.1641-59.

[10] Banerjee, A. and E. Duflo (2005), "Growth Theory through the Lens of Development Economics", *Handbook of Economic Growth*, vol. 1(A), P. Aghion and S. Durlauf eds., p. 473–554.

[11] Banerjee, A., and B. Moll (2010), "Why Does Misallocation Persist?", *American Economic Journal: Macroeconomics*, 2(1), p.189–206.

[12] Berger, A. and G. Udell (1995), "Relationship lending and lines of credit in small firm finance", *Journal of Business*, 68, p. 351-82.

[13] Bertrand, M., P. Mehta and S. Mullainathan (2002), "Ferreting out Tunnelling: An application to Indian business groups", *Quarterly Journal of Economic*s 117(1), p.121-48

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

[15] Bewley, T. (1977), "The permanent income hypothesis: A theoretical formulation", *Journal of Economic Theory* 16(2), p. 252-92.

[16] Bizer, D. and P. DeMarzo (1992), "Sequential Banking", *Journal of Political Economy* 100(1), p.41-61.

[17] Bond, P. (2004), "Bank and non-bank financial intermediation", *Journal of Finance*, 59(6), p. 2489-2529.

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

[19] Bond, S., J. 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), p.153-65.

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

[21] Buera, F., and Y. Shin (2013), "Financial Frictions and the Persistence of History: A Quantitative Exploration", *Journal of Political Economy* 121(2), p.221–72.

[22] Carbó-Valverde, S., F. Rodríguez-Fernández and G. Udell (2008), "Bank Lending, Financing Constraints and SME Investment", *Federal Reserve Bank of Chicago WP* 2008-04.

[23] Caspar, C., A. Dias and H-P. Elsrodt (2010), "The five attributes of enduring family businesses", McKinsey & Company.

[24] Cetorelli, N. (2001), “Competition Among Banks: Good or Bad?”, *Economic Perspectives* 25(2), Federal Reserve Bank of Chicago.

[25] Chava, S., and A. Purnanandam (2009), “The effect of banking crisis on bank-dependent borrowers”, Working paper.

[26] Chirinko, R. and J. Elston (2006), “Finance, control and profitability: the influence of German banks”, *Journal of Economic Behaviour and Organization*, 59, p. 69-88.

[27] Cleary, S. (1999), “The relationship between firm investment and financial status”, *Journal of Finance*, 54, p. 673-92.

[28] Clementi, G-L. and H. Hopenhayn (2006), “A Theory of Financing Constraints and Firm Dynamics”, *Quarterly Journal of Economics*, 121(1), p.229-65.

[29] Cole, H. and N. Kocherlakota (2001), “Efficient Allocations with Hidden Income and Hidden Storage”, *Review of Economic Studies* 68, p.523-42.

[30] Cooper, R. and J. Ejarque (2003), “Financial Frictions and Investment: A Requiem in Q”, *Review of Economic Dynamics*, 6, p. 710-28.

[31] Crespí, R. and A. Martín-Oliver (2015), “Do Family Firms have Better Access to External Finance during Crises?”, *Corporate Governance: An International Review* 23(3), p. 249–65.

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

[33] D’Espallier, B., J. Huybrechts and F. Iturriaga (2011), “Analyzing firm-varying investment-cash flow sensitivities and cash-cash flow sensitivities: A Bayesian approach”, *Spanish Journal of Finance and Accounting* 40, p.439-67.

[34] De Mel, S., D. McKenzie and C. Woodruff (2008), “Returns to Capital in Microenterprises: Evidence from a Field Experiment,” *Quarterly Journal of Economics* 123(4), p.1329-72.

[35] Degryse, H. and S. Ongena (2005), “Distance, Lending Relationships, and Competition”, *Journal of Finance* 60, p.2757-88.

[36] 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, p. 90-109.

[37] Diamond, D. (1984), “Financial intermediation and delegated monitoring”, *Review of Economic Studies*, 51, p. 393-414.

[38] Dubois, P., B. Julien and T. Magnac (2008), “Formal and Informal Risk Sharing in LDCs: Theory and Empirical Evidence,” *Econometrica* 76(4), p.679–725

[39] Elston, J. (1996), “Investment, liquidity constraints and bank relationships: evidence from German manufacturing firms”, Discussion Paper 1329, CEPR.

[40] Elyasiani, E. and Goldberg, L. (2004), “Relationship lending: a survey of the literature”, *Journal of Economics and Business* 56, p. 315-30.

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

[42] Evenson, R., and D. Gollin (2003), “Assessing the Impact of the Green Revolution, 1960 to 2000”, *Science* 300(5620), p.758–62.

[43] Fabbri, D. (2010), “Law Enforcement and Firm Financing: Theory and Evidence”, *Journal of the European Economic Association* 8(4), p.776-816.

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

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

[46] Fohlin, C. (1998), “Relationship banking, liquidity and investment in the German industrialization”, *Journal of Finance*, 53(5), p. 1737-58.

[47] Fohlin, C. and F. Iturriaga (2010), “Bank Relationships, Ownership Concentration, and Investment Patterns of Spanish Corporate Firms”, working paper, Johns Hopkins University.

[48] 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 658, European Central Bank.

[49] 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), p. 1941-73.

[50] 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, p. 1579-1603.

[51] Gibson, M., (1995), “Can bank health affect investment? Evidence from Japan”, *Journal of Business* 68, p.281–308.

[52] Giné, X., and R. Townsend (2004), “Evaluation of Financial Liberalization: A General Equilibrium Model with Constrained Occupation Choice”, *Journal of Development Economics*, 74(2), p.269–307.

[53] Gomes, J., (2001), “Financing Investment”, *American Economic Review*, 90(5), p. 1263-85.

[54] Green, D. and E. Liu (2015), “Growing Pains in Financial Development: Institutional Weakness and Investment Efficiency”, working paper, MIT.

[55] 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*, Banco de España.

[56] 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, p.67-88.

[57] Houston, J. and C. James (2001), "Do relationship have limits? Bank relationships, financial constraints and investment", *Journal of Business*, 74(3), p. 347-74.

[58] Hubbard, R. (1998), "Capital Market imperfections and investment", *Journal of Economic Literature*, 36, March, p. 193-225.

[59] Jaffee, D. and T. Russell (1976), "Imperfect information, uncertainty and credit rationing", *Quarterly Journal of Economics*, 91, p. 651-66.

[60] Jensen, M. and W. Meckling (1976), "Theory of the firm: managerial behaviour, agency costs, and ownership structure", *Journal of Financial Economics*, 1 p. 305-60.

[61] Jeong, H. and R. Townsend (2007), "Sources of TFP Growth: Occupational Choice and Financial Deepening", *Economic Theory* 32(1): p. 179–221.

[62] Jiménez, G., V. Salas and J. Saurina (2006), "Determinants of collateral", *Journal of Financial Economics*, 81, p. 255-81.

[63] Jiménez, G., J. Lopez and J. Saurina (2009), "Empirical analysis of corporate credit lines", *Review of Financial Studies* 22, p.5069-98.

[64] Jiménez, G., S. Ongena, J. Peydró and J. Saurina (2012), "Credit supply and Monetary Policy: Identifying bank balance-sheet channel with loan applications", *American Economic Review*, 102(5), 2301-26.

[65] Jiménez, G., S. Ongena, J. Peydró and J. Saurina (2014), "Hazardous Times for Monetary Policy: What Do Twenty-Three Million Bank Loans Say About the Effects of Monetary Policy on Credit Risk?", *Econometrica*, 82(2), 463-505.

[66] Jiménez, G., S. Ongena, J. Peydró and J. Saurina (2016), "Macroprudential Policy, Countercyclical Bank Capital Buffers, and Credit Supply: Evidence from the Spanish Dynamic Provisioning Experiment", *Journal of Political Economy*, forthcoming.

[67] Kang, J., and R. Stulz (2000), "Do banking shocks affect borrowing firm performance? An analysis of the japanese experience", *Journal of Business* 73, p.1–23.

[68] Kaplan, S. and Zingales, L. (1997), "Do investment-cash flow sensitivities provide useful measures of financing constraints?", *Quarterly Journal of Economics*, 112(1), p. 169-215.

[69] Kaplan, S. and Zingales, L. (2000), "Investment-cash flow sensitivities are not valid measures of financing constraints", *Quarterly Journal of Economics*, 115(2), p.707-12.

[70] Karaivanov, A. (2012), "Financial Constraints and Occupational Choice in Thai Villages", *Journal of Development Economics* 97(2), p.201-20.

[71] Karaivanov, A. and R. Townsend (2014), “Dynamic Financial Constraints: Distinguishing Mechanism Design from Exogenously Incomplete Regimes”, *Econometrica* 82(3), p.887-959.

[72] Khanna, T. and Y. Yafeh (2007), “Business Groups in emerging markets: paragons or parasites”, *Journal of Economic Literature*, vol. XLV, p. 331-72.

[73] Khwaja, A., and A. Mian (2008), “Tracing the impact of bank liquidity shocks: Evidence from an emerging market”, *American Economic Review* 98, p.1413–42.

[74] Kinnan, C. (2014), “Distinguishing Barriers to Insurance in Thai Villages”, working paper.

[75] Kinnan, C. and R. Townsend (2012), “Kinship and Financial Networks, Formal Financial Access and Risk Reduction,” *American Economic Review* 102.

[76] Kirpalani, R. (2016), “Endogenously Incomplete Markets with Equilibrium Default”, working paper.

[77] Ligon, E. (1998), “Risk Sharing and Information in Village Economies,” *Review of Economic Studies* 66(4), p.847-64.

[78] López-Gracia, J., and S. Sánchez-Andújar (2007), “Financial structure of the family business: Evidence from a group of small Spanish firms”, *Family Business Review* 20(4), p.269–87.

[79] Maury, B. (2006), ”Family ownership and firm performance: empirical evidence from Western European corporations”, *Journal of Corporate Finance* 12, p. 321-41.

[80] Mestieri, M., J. Schauer and R. Townsend (2016), “Human Capital Accumulation and Occupational Choice in the Process of Development”, Working paper.

[81] Moll, B. (2014), “Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?”, *American Economic Review* 104(10), p.3186-3221.

[82] Morck, R. (2005), *A History of Corporate Governance around the World*, Randal K. Morck ed., University of Chicago Press.

[83] Myers, S. and Majluf, N. (1984), “Corporate financing and investment decisions when firms have information that investors do not have”, *Journal of Financial Economics*, 13, p. 187-221.

[84] Paulson, A. and R. Townsend (2004), “Entrepreneurship and Financial Constraints in Thailand”, *Journal of Corporate Finance* 10(2), p.229-62.

[85] Paulson, A., R. Townsend and A. Karaivanov (2006), “Distinguishing limited liability from moral hazard in a model of entrepreneurship”, *Journal of Political Economy* 114(1), p.100-44.

[86] Petersen, M., and R. Rajan (1994), “The Benefits of Lending Relationships: Evidence from Small Business Data”, *Journal of Finance* 49, p. 3–37.

[87] Petersen, M., and R. Rajan (1995), “The effect of credit market competition on lending relationship”, *Quarterly Journal of Economics* 110, p. 407–43.

[88] Phelan, C. and R. Townsend (1991), “Computing Multi-Period, Information-Constrained Optima”, *Review of Economic Studies* 58(5), p.853-81.

[89] Pindado, J. and C. De la Torre (2011), “Family control and investment–cash flow sensitivity: Empirical evidence from the Euro zone”, *Journal of Corporate Finance* 17(5), p. 1389–1409.

[90] Prescott, E.C. and R. Townsend (1984), “Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard”, *Econometrica* 52, p. 21-45.

[91] Rogerson, W. (1985), “The First-Order Approach to Principal-Agent Problems”, *Econometrica* 53(6), p.1357-67

[92] Samphantharak, K. (2003), *Internal capital markets in business groups*, PhD Thesis, University of Chicago.

[93] Samphantharak, K. and R. Townsend (2016), “Risk and Return in Village Economies”, working paper, MIT.

[94] Schiantarelli, F. (1996), “Financial constraints and investment: Methodological issues and international evidence”, *Oxford Review of Economic Policy*, 12(2), p. 70-89.

[95] Schmid, L. (2009), “A Quantitative Dynamic Agency Model of Financing Constraints”, Working paper, Duke University.

[96] Schnabl, P. (2010), “Financial globalization and the transmission of bank liquidity shocks: Evidence from an emerging market”, Working Paper.

[97] Sraer, D. and D. Thesmar (2004), “Performance and behaviour of family firms: Evidence from the French stock market”, *CEPR Discussion Paper*.

[98] Stiglitz, J. and A. Weiss (1981), “Credit rationing in markets with imperfect information”, *American Economic Review*, 71, p.393-410.

[99] Strebulaev, I. and T. Whited (2012), “Dynamic Models and Structural Estimation in Corporate Finance”, *Foundations and Trends in Finance* 6, p.1-163.

[100] Townsend, R. (1979), “Optimal contracts and competitive markets with costly state verification”, *Journal of Economic Theory*, 21, p. 265-93.

[101] Townsend, R. and K. Ueda (2010), “Welfare Gains From Financial Liberalization”, *International Economic Review* 51(3), p. 553-97

[102] Udry, C. and S. Anagol (2006), “The Return to Capital in Ghana”, *American Economic Review* 96(2), p.388-93.

[103] Vickery, J. (2005), “Banking relationship and the credit cycle. Evidence from the Asian financial crisis”, Working Paper, Federal Reserve Bank of New York.

[104] Vuong, Q. (1989), "Likelihood ratio tests for model selection and non-nested hypotheses", *Econometrica*, 57(2), p.307-33.

[105] Williamson, S. (1987), "Costly monitoring, optimal contracts and equilibrium credit rationing", *Quarterly Journal of Economics*, 102, p. 135-45.

[106] Zhorin, V. and R. Townsend (2014), "Spatial Competition among Financial Service Providers and Optimal Contract Design", working paper, MIT.

# Appendix A – Computation

## A1. Linear Programs

We discretize the dynamic programs from Section 2.2 by assuming that all state and choice variables belong to finite grids. Dividends, $c$ belong to the grid $C$, capital ($k$ and $k'$) belongs to the grid $K$, effort to the grid $Z$, output to the grid $Q$ and promised utility to the grid $W$. In principle the grids can be arbitrarily fine, although in practice we are constrained by computer time and memory requirements.

Second, we use the discretized problems to define the joint probability distribution over the choice variables such as effort, $z$, capital $k'$, etc. and cash flow (net profits) $q$, conditional on the current state as new choice variables in a linear program. For example, acknowledging that $c$ is not an independent choice variable but can be expressed in terms of $k$ and $k'$, the joint probability distribution in the saving only regime can be written as $\pi(z, k', b', q, |k, b)$ where each $\pi(\tilde{z}, \tilde{k}', \tilde{b}', \tilde{q}, |\tilde{k}, \tilde{b})$ equals the joint probability of the allocation $(\tilde{z}, \tilde{k}', \tilde{b}', \tilde{q}) \in Z \times K \times B \times Q$ in the optimal contract for given capital and savings/debt state $\tilde{k} \in K$, $\tilde{b} \in B$.

The **saving only (S)** and the **borrowing and lending (B)** financial regimes solve the following linear dynamic program, for any $k \in K$, $b \in B$:

$$\begin{aligned}
\Pi(k,b) \quad = \quad & \max_{\pi(z,k',b',q|k,b)\geq 0} \sum_{z,k',b',q} \pi(z,k',b',q|k,b)[u(q - \hat{g}(k,k') + (1-\delta)k - k' + b' - Rb, z) + \beta\Pi(k',b')] \\
& \text{s.t. } \sum_{b'} \pi(\overline{z}, \overline{k}', b', \overline{q}|k,b) = P(\overline{q}|\overline{z}, \overline{k}') \sum_{b',q} \pi(\overline{z}, \overline{k}', b', q|k,b) \text{ for all } \overline{z}, \overline{k}', \overline{q} \in Z \times K \times Q \\
& \text{s.t. } \sum_{z,k',b',q} \pi(z,k',b',q|k,b) = 1.
\end{aligned}$$

where we used $c = q - g(k, k' - (1-\delta)k) + (1-\delta)k - k' + b' - Rb$ from the firm's budget constraint. The constraints on the probabilities $\pi(z, k', b', q|k, b)$ ensure that they are non-negative, add up to $1$, and conform with the net profits production technology represented by $P(q|z, k)$. In the saving only regime we set the upper bound on the grid $B$ to zero, corresponding to no borrowing. In the borrowing and lending regime the upper bound on $B$ is set to a positive value.

To compute solutions to the **moral hazard (MH)** regime we use the joint probability distribution, $\pi(z, c, k', w', q|k, w)$ over the possible allocations for effort $z$, capital $k'$, net profits, $q$, dividends $c$ and future discounted ("promised") utility $w'$, assuming that all variables belong to discrete grids. The grid bounds for the promised utility grid $W$ are $w_{\min}$ and $w_{\max}$ where, compatible with incentive provision, $w_{\min}$ is the discounted utility of having the minimum possible dividend, $c_{\min}$ and applying the lowest possible effort level, $z_{\min}$ forever and $w_{\max}$ equals the discounted utility of having the maximum possible dividend $c_{\max}$ and lowest possible effort level, $z_{\min}$ forever. See KT (2014) for more details.

We re-write the problem in (1) as a dynamic linear program in the probabilities $\pi$. With a large number of firms, $\pi$ will also correspond to the frequency distribution we see in the data, if all variables were observed. The objective function is:

$$V(k,w) = \max_{\pi(z,c,k',w',q|k,w)\geq 0} \sum_{z,c,k',w',q} \pi(z,c,k',w',q|k,w)[q-c-k'+(1-\delta)k+R^{-1}V(k',w')]$$

The incentive compatibility constraints (ICC) can be written, $\forall \bar{z}, \hat{z} \neq \bar{z}$ as

$$\sum_{c,k',w',q} \pi(\bar{z},c,k',w',q|k,w)[u\left(c-\hat{g}(k,k'),\bar{z}\right)+\beta w'] \geq$$
$$\geq \quad \sum_{c,k',w',q} \pi(\bar{z},c,k',w',q|k,w)\frac{\text{Prob}(q|\hat{z},k')}{\text{Prob}(q|\bar{z},k')}[u\left(c-\hat{g}(k,k'),\hat{z}\right)+\beta w']$$

The promised utility $w$ entering the period must be delivered, so for $\forall k$ we must have ("promise keeping"):

$$\sum_{z,c,k',w',q} \pi(z,c,k',w',q|k,w)[u\left(c-\hat{g}(k,k'),z\right)+\beta w'] = w$$

In addition, we also have the technological feasibility constraints:

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

and, finally, the "adding-up" constraint:

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

## A2. Grids, parameters, and functional forms

To compute and estimate the models with our yearly data we fix the firms' and intermediary's discount factors to $\beta = .95$ and $1/R = .95$. We calibrate the depreciation rate as $\delta = .18$, which corresponds to the median value of the yearly tangible assets depreciation computed from our data.[46]

We convert the data into model units (normalize) by dividing all currency values for the assets $k$, investment $I$ and cash flow $q$ by the 90-th percentile of the distribution of tangible fixed assets ($k$) in the data. In the baseline estimation runs, we use a five-point grid, $K$ for capital, corresponding to the 10th, 30th, 50th, 70th and 90th asset

[46] We estimated the empirical distribution function of firm-level depreciation rates (measured by the firms' accounting depreciation expressed as percentage of tangible fixed assets) for 1997-2007. For each year, we obtained the median depreciation rate across firms and we set the value of $\delta$ equal to the average of these median depreciation rates.

percentiles in the data.[47] For instance, for the 2004-07 panel, $K = \{0, .025, .084, .235, 1\}$. We construct the cash flow grid $Q$ following the same procedure, setting its five points to the corresponding percentiles from the data. For example, for the 2004-07 panel, we have $Q = \{.005, .013, .035, .087, .365\}$. The grid for investment $I$ used in the computation of the various models is obtained from the $K$ grid using all possible values that $k' - (1-\delta)k$ can take when $k, k' \in K$ and the grid $K$ is defined as explained earlier. Note that this allows for both upward and downward adjustments of the capital stock. The effort grid $Z$ is set to the three-point vector $\{.01, .505, 1\}$ which can be though of as 'low', 'medium' or 'high' effort levels.

We use the following functional form for the firms' (owners) payoff,

$$u(c, z) = c - \chi z^{\gamma}$$

where $\chi > 0$ is a parameter governing the relative cost of effort and $\gamma > 0$ parameterizes the curvature of the disutility of effort. In a robustness check we also allow for risk aversion by using $u(c, z) = \frac{c^{1-\sigma}}{1-\sigma} - \chi z^{\gamma}$ with $\sigma > 0$.

On the production side, remember that cash flow (net profits) $q$ is assumed to take on a finite number of values, $q_i$, $i = L, \ldots, \#Q$. We compute the probabilities $P(q|k, z)$ of obtaining cash flow level $q \in Q$ from capital level $k \in K$ and effort level $z \in Z$ using the data as follows. The grids $K$ and $Q$ are determined from the data percentiles as explained above. For each possible $q_i \in Q$ and $k_j \in K$, we then use a histogram function on the $q$ and $k$ data and take the frequency of observations with the given $q_i$ and $k_j$, call it $p(q_j|k_i)$. We then assign the probabilities $P(q|k, z)$ by distributing the observed frequency in the data, $p(q_j|k_i)$ for each $k_i \in K$ over the unobserved effort levels $z \in Z$.[48] The chosen specification provides a parsimonious way of calibrating the production technology to the available data, subject to the unobservability of the variable $z$.[49]

The assumed functional forms are not essential for our computational and estimation methods. Our linear programming approach is completely general and can use any (including non-convex) functional forms for the preferences and technology. The main reason we chose the specific forms used here is their parsimonious use of parameters which is important for computation speed.

[47] 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 added computation time.

[48] To do so we use the following matrix:

$$T = \begin{pmatrix} .7 & .5 & .25 & .1 & .05 \\ .25 & .4 & .5 & .4 & .25 \\ .05 & .1 & .25 & .5 & .7 \end{pmatrix}$$

In the matrix above the rows correspond to the three effort $(z)$ levels and the columns correspond to the five net profits $(q)$ levels. For example, the first column means that, for any $k \in K$, we assign .7 of the probability (obtained from the data), $p(q_1, k)$ that $q = q_1$ (the lowest net profit is realized) to $z = z_1$ (the lowest effort level), .25 of the probability to the medium effort level, $z = z_2$ and .05 to the highest effort level, $z = z_3$. The opposite order of these weights is used for $q = q_5$, the highest net profit level.

[49] In the estimation runs with capital adjustment costs we use a standard quadratic form from the literature (e.g., Bond et al., 2005),

$$g(k_t, I_t) = \frac{1}{2} b k_t \left[\frac{I_t}{k_t}\right]^2$$

where $b$ is a parameter to be estimated.

## Appendix B – Variables definitions

**Firm characteristics**

– *Age*: relative to the firm registration date.

– *Industry*: three- and four-digit CNAE-93 classification. Using the industry information we exclude financial firms.

**Firm-level economic variables**

– *Capital* ($k_t$): firm capital stock in year $t$ is measured by the book value of the firm's tangible fixed assets.

– *(Gross) investment* ($I_t$): defined as the sum of the absolute increase in the stock of tangible fixed assets between years $t$ and $t-1$ and the depreciation in year $t$.

– *Cash flow* ($q_t$)*:* defined as the sum of operating profits/losses and depreciation.

– *Sales* ($S_t$): measured using total sales revenue (not used in the structural estimation, only in summary statistics).

– *Total assets* ($A_t$): from the firms' balance sheets (not used in the structural estimation, only in summary statistics)

**Financial ratios**

– *Debt-to-assets* (leverage) ratio: total debt divided by total assets

– *Liquidity* ratio: cash to short term debt ratio.

**Firm bank relations**

– *Number of bank relations*: number of different Spanish credit institutions granting a loan to the firm. According to the current number of bank relations, we distinguish year-by-year three types of firms: unbanked, in single bank relationship (single-banked), and in multiple bank relationships (multi-banked). Empirically, to analyze differences across groups of firms defined according to their banking status, we consider firms that maintain continuously the same status over all the years they are present in the data (see Section 4)

## Appendix C – The likelihood

We have panel data, $\{\hat{k}_{jt}, \hat{I}_{jt}, \hat{q}_{jt}\}$ for $j = 1, ..., n$ and $t = 0, ..., T$, where the subscripts $j$ and $t$ denote firm and time respectively; $k$ is capital, $I$ is investment, and $q$ is cash flow. The data are assumed i.i.d. across firms. For each firm $j$, call $\hat{y}_j$ the vector of variables used in the estimation. In the different MLE runs we do $\hat{y}_j$ can consist of either cross-sectional variables, for example, period $t$ capital, investment and cash flow (that is $\hat{y}_j = (\hat{k}_{jt}, \hat{I}_{jt}, \hat{q}_{jt})$) or variables from different time periods, for example, firm investment and cash flow from dates $t$ and $t+1$, that is, $\hat{y}_j = (\hat{I}_{jt}, \hat{q}_{jt}, \hat{I}_{jt+1}, \hat{q}_{jt+1})$).

For given: (i) structural parameters of the model (preferences and technology), (ii) distribution parameters for unobserved heterogeneity (the initial distribution of debt/savings $b$ or promised utility, $w$), and (iii) a distribution parameter for the measurement error, we construct the likelihood, $\Lambda^m(\phi)$ of each model $m \in \{A, B, MH\}$ with respect to the data $\{\hat{y}_j\}_{j=1}^n$ where $\phi$ is the vector of all estimated parameters.

Following Karaivanov and Townsend (KT, 2014), call $s^1$ the observed state variable (capital, $k$) and $s^2$ the unobserved state variable (debt/savings, $b$ or promised utility, $w$) in the model. For given structural parameters $\phi^s$ for preferences and technology, the solution of each LP problem in Section 2 is a discrete joint probability distribution, $\pi(.|s^1, s^2)$, using which we can obtain the theoretical joint distribution in model regime $m$ over any vector of variables $y$ used in the estimation. Call this distribution $g^m(y|s^1, s^2; \phi^s)$.

We allow for the possibility that the data $\hat{y}$ contain additive Normally distributed measurement error with variance parameter $\gamma_{me}$ which we estimate.[50] Next, we need to initialize the states $s^1$ and $s^2$. The observed state $s^1$ (capital $k$) is initialized using the actual data $\{\hat{k}_j\}_{j=1}^n$ from the (initial) period used in the estimation, discretized over the grid $K$ via histogram function (see KT, 2014). Call the resulting discrete probability distribution $h(s_0^1)$. The initial distribution of the unobserved states $s_0^2$ is treated as unobservable heterogeneity and parametrized by $\phi^d$. [51] Overall, we obtain the initial joint state distribution $h(s_0^1, s_0^2; \phi^d)$.

To write the likelihood of model $m$ with the given observations on variables $y$, construct

$$f^m(y; \phi^s, \phi^d) \equiv \sum_{s_0^2} g^m(y|s_0^1, s_0^2; \phi^s) h(s_0^1, s_0^2; \phi^d),$$

which is the joint distribution of the observable variables $y$ given parameters $\phi^s, \phi^d$ integrated over the distribution of the unobserved state.

Let $\Phi\left(.|\mu, \sigma\right)$ be the pdf of a Normal distribution with mean $\mu$ and variance $\sigma^2$. Given the assumed form of measurement error, the likelihood of observing data $\hat{y}_j$ (for instance $\hat{I}_j, \hat{q}_j$) relative to grid point $y_r \in Y$ for any $r = 1, ..., \#Y$ is,

$$\prod_{l=1}^{L} \Phi\left(\hat{y}_j^l | y_r^l, \sigma^l(\gamma_{me})\right) \tag{5}$$

where $l = 1, ..., L$ indexes the distinct variables and/or time periods used in the estimation (e.g., $L = 2$ if $y = (I_t, q_t)$ and $L = 4$ if $y = (I_t, q_t, I_{t+1}, q_{t+1})$) and where $\sigma^l(\gamma_{me})$ is as defined in footnote 28.

Putting all the pieces together, the likelihood of the data $\{\hat{y}_j\}_{j=1}^n$ in model $m$ given initial state distribution $h(s_0^1, s_0^2; \phi^d)$, at parameters $\phi \equiv (\phi^s, \phi^d, \gamma_{me})$ is,

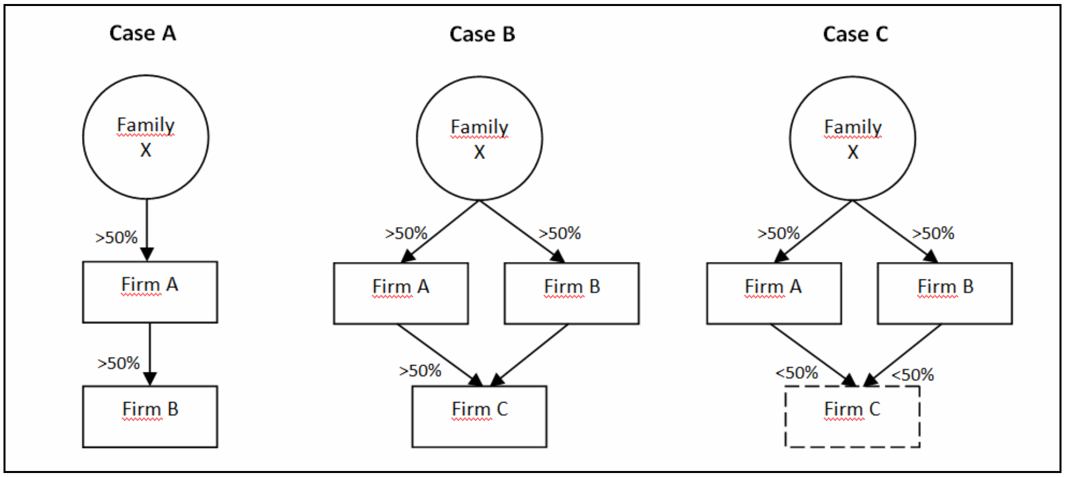
$$\Lambda^m\left(\phi\right) \equiv \sum_{j=1}^{n} \ln[\sum_{r=1}^{\#Y} f^m(y_r; \phi^s, \phi^d) \prod_{l=1}^{L} \Phi\left(\hat{y}_j^l | y_r^l, \sigma^l(\gamma_{me})\right)]. \tag{6}$$

We maximize the log-likelihood $\Lambda^m(\phi)$ by choice of the parameters $\phi$ for each of the three financial regimes (A, B, and MH). We first run an extensive grid search over the parameters and then use a global optimization method starting from the best-fitting candidate parameters from the grid search. Standard errors are computed via bootstrapping, repeatedly drawing with replacement from the data up to the original sample size. Finally, we follow Vuong (1989) and use the maximized likelihood for each model (A, B and MH) to compute an asymptotic test statistic and formally compare (pairwise) across the alternative financial regimes.

[50]Specifically, we assume that the measurement error of variable $x$ has distribution $N(0, (\gamma_{me}\chi(x))^2)$ where $\chi(x)$ denotes the variable's grid range, $\chi(x) \equiv x_{\max} - x_{\min}$. More complex versions of measurement error parametrized by additional parameters can be considered at the cost of added computing time.

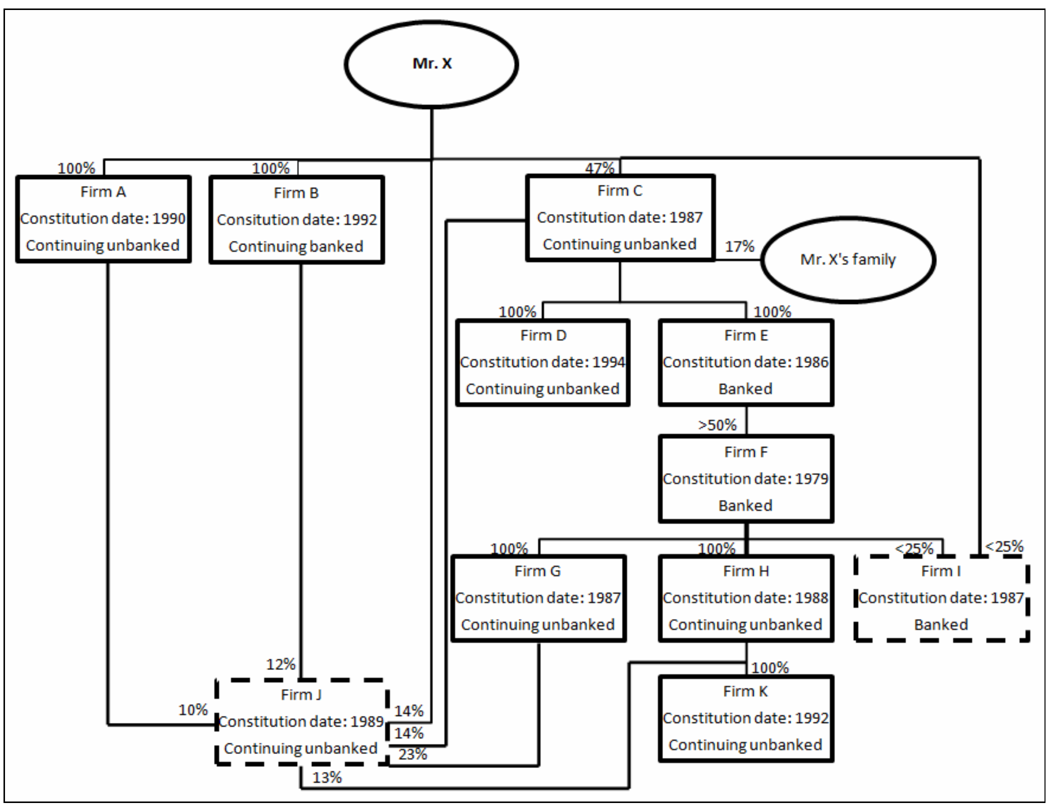
[51]In our baseline estimation runs we assume that the initial distributions of the state variables $s^1$ and $s^2$ are independent and that the unobserved state is distributed $\Omega(s_0^2, \phi^d) = N(\mu_{s^2}, \gamma_{s^2}^2)$. This assumption is not essential, see KT (2014).

**Figure 2. Examples of family networks**



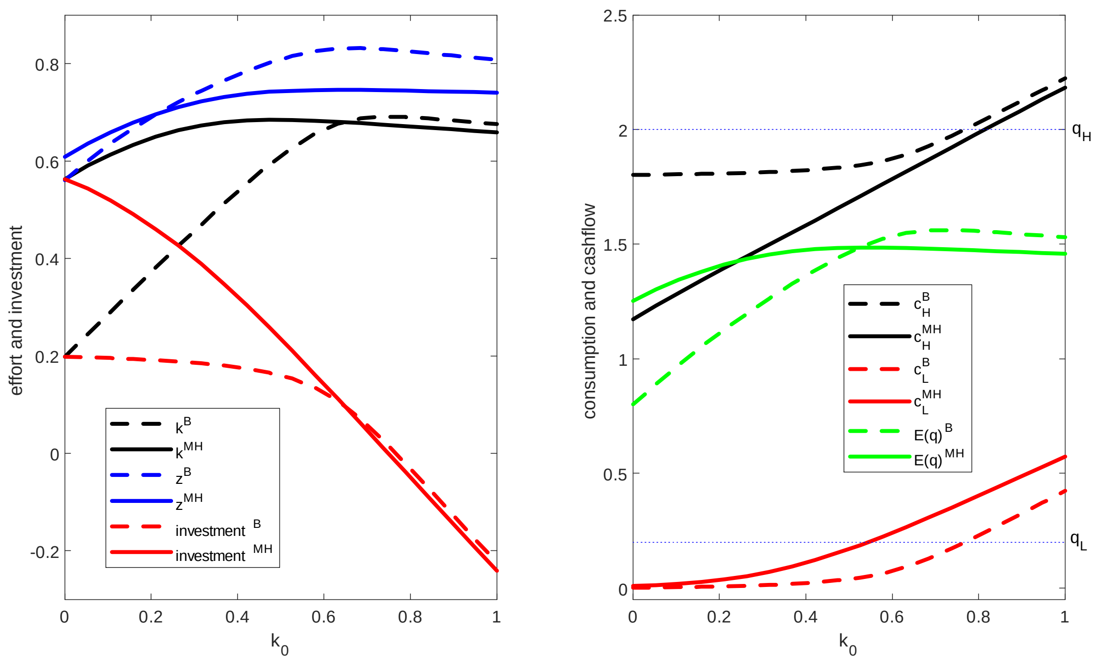

Note: Solid lines represent firms included in the sample as network firms and dashed lines represent firms not identified as network firms.

**Figure 3. Case study of a family network**




Note: Solid lines represent firms included in the sample as network firms and dashed lines represent firms not identified as network firms.

Figure 5: Computed solutions with risk aversion



# Table 1 - Sample comparisons

| | (1) | (2) | (3) |
|---|---|---|---|
| **A. all years, 1997-2007** | | | |
| observations | 3718286 | 1493181 | |
| number of firms | 805118 | 385539 | |
| mean age | 11 | 11.6 | |
| mean assets, k | 991 | 1446 | |
| median k | 79 | 105 | |
| stdev k | 29796 | 41723 | |
| mean cash flow, q | 303 | 501 | |
| median q | 35 | 45 | |
| stdev q | 10178 | 15686 | |
| mean investment, I | 196 | 273 | |
| median I | 10 | 12 | |
| stdev I | 14975 | 7218 | |
| percent family firms | 32% | 34% | |
| percent unbanked | 21% | 24% | |
| percent single-banked | 31% | 18% | |
| percent multi-banked | 48% | 58% | |
| | | | |
| **B. 2004-2007** | | | |
| observations | 1782431 | 711985 | 402980 |
| number of firms | 653804 | 290276 | 100745 |
| mean age | 11.4 | 11.7 | 14.5 |
| mean assets, k | 1070 | 1455 | 2073 |
| median k | 81 | 95 | 175 |
| stdev k | 32897 | 43627 | 55548 |
| mean cash flow, q | 321 | 510 | 751 |
| median q | 33 | 37 | 73 |
| stdev q | 11670 | 17918 | 20527 |
| mean investment, I | 206 | 272 | 380 |
| median I | 9 | 9 | 20 |
| stdev I | 20548 | 7168 | 8539 |
| percent family firms | 32% | 33% | 42% |
| percent unbanked | 21% | 27% | 17% |
| percent single-banked | 33% | 22% | 16% |
| percent multi-banked | 46% | 51% | 67% |

Notes:

(1) all observations with non-missing (k,I,q) data from any of the listed years

(2) observations of firms with non-missing (k,I,q) data from any of the listed years and continuing bank status

(3) observations of firms with non-missing (k,I,q) data present in all listed years and continuing bank status (balanced panel)

* All means, medians, standard deviations (stdev) and percentages are computed at the observation level.

## Table 2A - Summary statistics by banking and family status - all balanced panels

| | unbanked firms | | | single-banked firms | | | multi-banked firms | | |
|---|---|---|---|---|---|---|---|---|---|
| | all | family | non family | all | family | non family | all | family | non family |
| observations | 128147 | 22005 | 106142 | 107031 | 28271 | 78760 | 645549 | 326177 | 319372 |
| number of firms | 26601 | 4097 | 22504 | 21150 | 4884 | 16266 | 87700 | 40893 | 46807 |
| percent of all | | 17.2% | 82.8% | | 26.4% | 73.6% | | 50.5% | 49.5% |
| *Data used in the MLE* | | | | | | | | | |
| mean assets (k) | 415 | 246 | 450 | 456 | 353 | 494 | 2556 | 1004 | 4141 |
| median k | 20 | 24 | 19 | 77 | 81 | 76 | 266 | 261 | 272 |
| stdev k | 4376 | 1019 | 4785 | 3077 | 1238 | 3509 | 58115 | 6592 | 82324 |
| mean cash flow (q) | 80 | 144 | 66 | 109 | 108 | 110 | 954 | 409 | 1511 |
| median q | 10 | 16 | 9 | 28 | 40 | 25 | 126 | 132 | 117 |
| stdev q | 2953 | 6979 | 657 | 657 | 371 | 733 | 20631 | 2113 | 29243 |
| mean investment (I) | 47 | 33 | 49 | 65 | 55 | 68 | 497 | 226 | 774 |
| median I | 1 | 1 | 1 | 5 | 8 | 5 | 45 | 46 | 43 |
| stdev I | 1191 | 376 | 1297 | 866 | 476 | 968 | 9368 | 2337 | 13101 |
| *Financial indicators and others* | | | | | | | | | |
| mean total assets (A) | 2095 | 1262 | 2268 | 2256 | 1160 | 2650 | 10235 | 4524 | 16068 |
| median A | 255 | 283 | 249 | 389 | 494 | 355 | 1477 | 1466 | 1494 |
| stdev A | 40496 | 10399 | 44242 | 78537 | 2879 | 91534 | 128585 | 21957 | 181276 |
| mean sales revenue (S) | 697 | 878 | 656 | 1068 | 1315 | 975 | 10802 | 5038 | 16745 |
| median S | 208 | 320 | 190 | 416 | 699 | 351 | 1941 | 2122 | 1660 |
| stdev S | 5278 | 7623 | 4587 | 4868 | 2135 | 5552 | 140452 | 22027 | 198737 |
| mean q/A ratio | 0.08 | 0.11 | 0.08 | 0.10 | 0.11 | 0.10 | 0.11 | 0.11 | 0.10 |
| median q/A | 0.06 | 0.08 | 0.05 | 0.09 | 0.09 | 0.08 | 0.09 | 0.10 | 0.09 |
| stdev q/A | 1.31 | 0.18 | 1.44 | 0.15 | 0.14 | 0.16 | 0.12 | 0.12 | 0.11 |
| mean q/k ratio | 2.90 | 3.46 | 2.78 | 3.14 | 2.45 | 3.40 | 2.53 | 2.03 | 3.04 |
| median q/k | 0.36 | 0.54 | 0.32 | 0.33 | 0.44 | 0.29 | 0.49 | 0.51 | 0.47 |
| stdev q/k | 54 | 47 | 55 | 56 | 27 | 64 | 73 | 20 | 102 |
| mean debt/assets ratio (D/A) | 0.38 | 0.42 | 0.37 | 0.60 | 0.60 | 0.60 | 0.70 | 0.71 | 0.69 |
| median D/A | 0.32 | 0.40 | 0.30 | 0.64 | 0.63 | 0.64 | 0.74 | 0.74 | 0.73 |
| stdev D/A | 0.31 | 0.30 | 0.31 | 0.26 | 0.25 | 0.26 | 0.19 | 0.18 | 0.20 |
| mean liquidity ratio (LR) | 3.36 | 2.17 | 3.62 | 0.95 | 0.71 | 1.03 | 0.22 | 0.18 | 0.25 |
| median LR | 0.54 | 0.55 | 0.54 | 0.21 | 0.22 | 0.20 | 0.07 | 0.07 | 0.06 |
| stdev LR | 49.6 | 24.5 | 53.6 | 9.0 | 3.4 | 10.2 | 2.1 | 1.1 | 2.7 |
| mean age | 11.6 | 11.7 | 11.6 | 10.7 | 11.2 | 10.6 | 15.1 | 14.7 | 15.5 |

* All means, medians, standard deviations (stdev) and percentages are computed at the observation level.

## Table 2B - Summary statistics by family firm subcategories - all balanced panels

| | unbanked | | | single-banked | | | multi-banked | | |
|---|---|---|---|---|---|---|---|---|---|
| | pure family | fam-netw. | non family | pure family | fam-netw. | non family | pure family | fam-netw. | non family |
| observations | 19390 | 2615 | 106142 | 26703 | 1568 | 78760 | 302971 | 23206 | 319372 |
| number of firms | 3609 | 491 | 22504 | 4603 | 285 | 16266 | 38333 | 2922 | 46807 |
| percent of all | 13.6% | 1.8% | 84.6% | 21.8% | 1.3% | 76.9% | 43.7% | 3.3% | 53.4% |
| *Data used in the MLE* | | | | | | | | | |
| mean assets (k) | 221 | 429 | 450 | 300 | 1255 | 494 | 726 | 4632 | 4141 |
| median k | 23 | 31 | 19 | 78 | 153 | 76 | 244 | 854 | 272 |
| stdev k | 936 | 1487 | 4785 | 884 | 3667 | 3509 | 3368 | 21177 | 82324 |
| mean cash flow (q) | 57 | 786 | 66 | 96 | 305 | 110 | 299 | 1854 | 1511 |
| median q | 16 | 21 | 9 | 38 | 88 | 25 | 124 | 437 | 117 |
| stdev q | 196 | 20230 | 657 | 220 | 1270 | 733 | 917 | 7038 | 29243 |
| mean investment (I) | 30 | 61 | 49 | 49 | 154 | 68 | 175 | 895 | 774 |
| median I | 2 | 1 | 1 | 8 | 9 | 5 | 43 | 145 | 43 |
| stdev I | 335 | 600 | 1297 | 351 | 1409 | 968 | 2093 | 4370 | 13101 |
| *Financial indicators and others* | | | | | | | | | |
| mean total assets (A) | 861 | 4232 | 2268 | 1005 | 3800 | 2650 | 3239 | 21305 | 16068 |
| median A | 257 | 561 | 249 | 476 | 1174 | 355 | 1363 | 5778 | 1494 |
| stdev A | 3765 | 28198 | 44242 | 2123 | 8082 | 91534 | 10866 | 70230 | 181276 |
| mean sales revenue (S) | 701 | 2342 | 656 | 1272 | 2088 | 975 | 3945 | 19366 | 16745 |
| median S | 315 | 397 | 190 | 692 | 854 | 351 | 2011 | 5886 | 1660 |
| stdev S | 1114 | 22934 | 4587 | 1939 | 4275 | 5552 | 8574 | 75228 | 198737 |
| mean q/A ratio | 0.11 | 0.09 | 0.08 | 0.11 | 0.11 | 0.10 | 0.11 | 0.10 | 0.10 |
| median q/A | 0.08 | 0.04 | 0.05 | 0.09 | 0.08 | 0.08 | 0.10 | 0.09 | 0.09 |
| stdev q/A | 0.18 | 0.21 | 1.44 | 0.13 | 0.15 | 0.16 | 0.12 | 0.09 | 0.11 |
| mean q/k ratio | 2.65 | 10.57 | 2.78 | 2.17 | 7.64 | 3.40 | 1.84 | 4.58 | 3.04 |
| median q/k | 0.56 | 0.39 | 0.32 | 0.44 | 0.39 | 0.29 | 0.51 | 0.51 | 0.47 |
| stdev q/k | 15 | 140 | 55 | 14 | 101 | 64 | 14 | 55 | 102 |
| mean debt/assets ratio (D/A) | 0.44 | 0.33 | 0.37 | 0.61 | 0.52 | 0.60 | 0.72 | 0.65 | 0.69 |
| median D/A | 0.42 | 0.25 | 0.30 | 0.64 | 0.52 | 0.64 | 0.75 | 0.69 | 0.73 |
| stdev D/A | 0.30 | 0.30 | 0.31 | 0.25 | 0.27 | 0.26 | 0.18 | 0.20 | 0.20 |
| mean liquidity ratio (LR) | 1.96 | 3.81 | 3.62 | 0.68 | 1.12 | 1.03 | 0.19 | 0.14 | 0.25 |
| median LR | 0.57 | 0.33 | 0.54 | 0.22 | 0.20 | 0.20 | 0.07 | 0.05 | 0.06 |
| stdev LR | 10.2 | 66.9 | 53.6 | 3.3 | 5.2 | 10.2 | 1.2 | 0.6 | 2.7 |
| mean age | 11.5 | 13.1 | 11.6 | 11.1 | 12.5 | 10.6 | 14.4 | 18.6 | 15.5 |

* All means, medians, standard deviations (stdev) and percentages are computed at the observation level.

## Table 2C - Summary statistics by industry - all balanced panels

| | unbanked firms | | | single-banked firms | | | multi-banked firms | | |
|---|---|---|---|---|---|---|---|---|---|
| | all | constr | non-constr | all | constr | non-constr | all | constr | non-constr |
| observations | 128147 | 39548 | 88599 | 107031 | 28718 | 78313 | 645549 | 129303 | 516246 |
| percent of all | | 30.9% | 69.1% | | 26.8% | 73.2% | | 20.0% | 80.0% |
| percent family firms | 17.2% | 12.9% | 19.1% | 26.4% | 19.6% | 28.9% | 50.5% | 48.4% | 51.1% |
| *data used in the MLE* | | | | | | | | | |
| mean assets (k) | 415 | 720 | 278 | 456 | 696 | 369 | 2556 | 1694 | 2772 |
| median k | 20 | 87 | 15 | 77 | 109 | 71 | 266 | 219 | 277 |
| stdev k | 4376 | 4197 | 4447 | 3077 | 2805 | 3167 | 58115 | 17334 | 64403 |
| mean cash flow (q) | 80 | 58 | 89 | 109 | 131 | 101 | 954 | 711 | 1015 |
| median q | 10 | 9 | 10 | 28 | 28 | 28 | 126 | 124 | 126 |
| stdev q | 2953 | 376 | 3543 | 657 | 818 | 586 | 20631 | 4817 | 22943 |
| mean investment (I) | 47 | 52 | 44 | 65 | 73 | 62 | 497 | 309 | 544 |
| median I | 1 | 0 | 1 | 5 | 2 | 7 | 45 | 32 | 48 |
| stdev I | 1191 | 1239 | 1169 | 866 | 962 | 828 | 9368 | 5302 | 10133 |
| mean age | 11.6 | 12.2 | 11.4 | 10.7 | 10 | 11 | 15.1 | 13.4 | 15.6 |

## Table 2D - Summary statistics by region - all balanced panels

| | unbanked firms | | | single-banked firms | | | multi-banked firms | | |
|---|---|---|---|---|---|---|---|---|---|
| | all | region 1 | region 2 | all | region 1 | region 2 | all | region 1 | region 2 |
| observations | 128147 | 98088 | 30059 | 107031 | 75362 | 31669 | 645549 | 489010 | 156539 |
| percent of all | | 76.5% | 23.5% | | 70.4% | 29.6% | | 75.8% | 24.2% |
| percent family | 17.2% | 16.5% | 19.5% | 26.4% | 25.7% | 28.2% | 50.5% | 49.3% | 54.4% |
| *data used in the MLE* | | | | | | | | | |
| mean assets (k) | 415 | 473 | 224 | 456 | 487 | 385 | 2556 | 2860 | 1608 |
| median k | 20 | 21 | 17 | 77 | 79 | 74 | 266 | 256 | 293 |
| stdev k | 4376 | 4932 | 1490 | 3077 | 3380 | 2193 | 58115 | 66270 | 14399 |
| mean cash flow (q) | 80 | 91 | 42 | 109 | 118 | 88 | 954 | 1090 | 532 |
| median q | 10 | 10 | 9 | 28 | 30 | 25 | 126 | 130 | 114 |
| stdev q | 2953 | 3370 | 338 | 657 | 720 | 473 | 20631 | 23566 | 4489 |
| mean investment (I) | 47 | 53 | 27 | 65 | 70 | 52 | 497 | 551 | 327 |
| median I | 1 | 1 | 1 | 5 | 6 | 4 | 45 | 45 | 43 |
| stdev I | 1191 | 1349 | 326 | 866 | 945 | 639 | 9368 | 10498 | 4194 |
| mean age | 11.6 | 11.9 | 10.8 | 10.7 | 10.9 | 10.3 | 15.1 | 15.5 | 14.1 |

* All means, medians, standard deviations (stdev) and percentages are computed at the observation level.

**Table 2E - Firm distribution by banking status, region, industry, size/age and ownership**

| Stratification | unbanked | single-banked | multi-banked |
|---|---|---|---|
| all | 15% | 12% | 73% |
| region 1 | 15% | 11% | 74% |
| region 2 | 14% | 14% | 72% |
| construction | 20% | 15% | 65% |
| no construction | 13% | 11% | 76% |
| small and young | 24% | 19% | 57% |
| large and old | 5% | 5% | 90% |
| family | 6% | 8% | 86% |
| non-family | 21% | 16% | 63% |

*Note: based on pooled data from all three balanced panels

## Table 4 - Vuong Test Model Comparisons - Firm Size and Age

| | A. whole sample | | | | B. small and young firms | | | | C. large and old firms | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Comparison: | MH v B | MH v S | B v S | Best fit | MH v B | MH v S | B v S | Best fit | MH v B | MH v S | B v S | Best fit |
| *1. 2004-07 data; investment (I) three-year panel* | | | | | | | | | | | | |
| 1.1 2004-06, unbanked | tie | S*** | S*** | **S** | tie | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| 1.2 2004-06, single-banked | B*** | S*** | S*** | **S** | MH*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| 1.3 2004-06, multi-banked | MH*** | MH*** | B*** | **MH** | B*** | S*** | S*** | **S** | MH*** | MH*** | S*** | **MH** |
| *2. 2004-07 data; investment and cashflow (I,q) two-year panels* | | | | | | | | | | | | |
| 2.1 2004,05 unbanked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** | MH*** | MH*** | B*** | **MH** |
| 2.2 2004,05 single-banked | B*** | S*** | B** | **B** | B*** | S*** | S*** | **S** | MH*** | MH*** | B*** | **MH** |
| 2.3 2004,05 multi-banked | MH*** | MH*** | S*** | **MH** | B*** | S*** | tie | **B,S** | MH*** | MH*** | tie | **MH** |
| 2.4 2004,07 unbanked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** | MH*** | MH*** | B* | **MH** |
| 2.5 2004,07 single-banked | B*** | tie | B*** | **B** | B*** | S*** | S*** | **S** | MH*** | MH*** | B*** | **MH** |
| 2.6 2004,07 multi-banked | MH*** | MH*** | B*** | **MH** | B*** | tie | B*** | **B** | MH*** | MH*** | B*** | **MH** |
| *3. 2004-07 data; assets, investment and cashflow (k,I,q) cross-sections* | | | | | | | | | | | | |
| 3.1 2007, unbanked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** | MH*** | MH*** | B** | **MH** |
| 3.2 2007, single-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** | tie | MH*** | B*** | **MH,B** |
| 3.3 2007, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |
| 3.4 2004, unbanked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** | MH*** | MH*** | B*** | **MH** |
| 3.5 2004, single-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** | MH* | MH*** | B*** | **MH** |
| 3.6 2004, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** | B*** | S*** | B*** | **B** |
| *4. 2001-03 data; assets, investment and cashflow (k,I,q) cross-sections* | | | | | | | | | | | | |
| 4.1 2002, unbanked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** | MH*** | MH** | S*** | **MH** |
| 4.2 2002, single-banked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** | B*** | S*** | B*** | **B** |
| 4.3 2002, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** | B*** | S*** | B*** | **B** |
| *5. 1997-00 data; assets, investment and cashflow (k,I,q) cross-sections* | | | | | | | | | | | | |
| 5.1 2000, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** | B*** | S*** | B*** | **B** |
| 5.2 2000, single-banked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** | B*** | S* | B*** | **B** |
| 5.3 2000, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |
| 5.4 1997, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | B* | **B** | B*** | S*** | B*** | **B** |
| 5.5 1997, single-banked | B*** | S*** | B*** | **B** | B*** | S*** | tie | **S,B** | B*** | S*** | B*** | **B** |
| 5.6 1997, multi-banked | B*** | S*** | B*** | **B** | tie | tie | tie | **S,B,MH** | B*** | S*** | B*** | **B** |

Notes: The listed regime is best fitting, including ties. ***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively.

## Table 5A - Vuong test model comparisons - Family status

| | A. family firms | | | | B. non-family firms | | | |
|---|---|---|---|---|---|---|---|---|
| Comparison: | MH v B | MH v S | B v S | Best fit | MH v B | MH v S | B v S | Best fit |
| *I. 2004-07 balanced panel, continuing banking status* | | | | | | | | |
| *1. investment (I) three-year panel* | | | | | | | | |
| 1.1 04-06, unbanked | B*** | S*** | S*** | **S** | MH*** | S*** | S*** | **S** |
| 1.2 04-06, single-banked | MH*** | MH*** | tie | **MH** | B*** | S*** | S*** | **S** |
| 1.3 04-06, multi-banked | tie | MH*** | B*** | **B,MH** | MH*** | MH*** | B*** | **MH** |
| *2. investment and cashflow (I,q) two-year panels* | | | | | | | | |
| 2.1 04,05 unbanked | B*** | S*** | S** | **S** | B*** | S*** | S*** | **S** |
| 2.2 04,05 single-banked | B*** | S*** | B*** | **B** | tie | S*** | S*** | **S** |
| 2.3 04,05 multi-banked | MH*** | MH*** | B*** | **MH** | MH*** | MH*** | B*** | **MH** |
| 2.4 04,07 unbanked | B*** | S*** | tie | **B,S** | B*** | S*** | S*** | **S** |
| 2.5 04,07 single-banked | B*** | MH*** | B*** | **B** | tie | S*** | S*** | **S** |
| 2.6 04,07 multi-banked | MH** | MH*** | B*** | **MH** | MH*** | MH*** | S*** | **MH** |
| *3. assets, investment and cashflow (k,I,q) cross-sections* | | | | | | | | |
| 3.1 2007, unbanked | B*** | S*** | B*** | **B** | MH*** | S*** | S*** | **S** |
| 3.2 2007, single-banked | B*** | S*** | B*** | **B** | B*** | S*** | tie | **S,B** |
| 3.3 2007, multi-banked | B*** | S*** | B*** | **B** | B*** | MH*** | B*** | **B** |
| 3.4 2004, unbanked | B*** | S*** | B* | **B** | B*** | S*** | B*** | **B** |
| 3.5 2004, single-banked | MH*** | MH*** | S*** | **MH** | B*** | S*** | B*** | **B** |
| 3.6 2004, multi-banked | tie | MH*** | B*** | **B,MH** | B*** | MH*** | B*** | **B** |
| *II. 2001-03 balanced panel, continuing banking status* | | | | | | | | |
| *4. investment (I) three-year panel* | | | | | | | | |
| 4.1 01-03, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |
| 4.2 01-03, single-banked | B*** | tie | B*** | **B** | B*** | S*** | S*** | **S** |
| 4.3 01-03, multi-banked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| *5. assets, investment and cashflow (k,I,q) cross-sections* | | | | | | | | |
| 5.1 2002, unbanked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| 5.2 2002, single-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 5.3 2002, multi-banked | B*** | S*** | B*** | **B** | B** | S*** | S*** | **S** |
| *III. 1997-00 balanced panel, continuing banking status* | | | | | | | | |
| *6. assets, investment and cashflow (k,I,q) cross-sections* | | | | | | | | |
| 6.1 2000, unbanked | tie | tie | B*** | **B,MH** | tie | S*** | S*** | **S** |
| 6.2 2000, single-banked | B*** | S*** | B** | **B** | B*** | S*** | S*** | **S** |
| 6.3 2000, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 6.4 1997, unbanked | B*** | S*** | B*** | **B** | MH** | S*** | S*** | **S** |
| 6.5 1997, single-banked | B*** | S*** | B*** | **B** | B** | S*** | S** | **S** |
| 6.6 1997, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |

Notes: The listed regime(s) is best fitting. ***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively.

## Table 5B - Vuong test model comparisons - Family status by size and age

| Comparison: | A. family firms MH v B | MH v S | B v S | Best fit | B. non-family firms MH v B | MH v S | B v S | Best fit |
|---|---|---|---|---|---|---|---|---|
| *IV. Small and young firms, continuing banking status; (k,I,q) cross-sections* | | | | | | | | |
| 7.1 2007, unbanked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| 7.2 2007, single-banked | B*** | S** | B*** | **B** | B*** | S*** | S*** | **S** |
| 7.3 2007, multi-banked | B*** | tie | B*** | **B** | B*** | S*** | B*** | **B** |
| 7.4 2004, unbanked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| 7.5 2004, single-banked | tie | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| 7.6 2004, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 8.1 2002, unbanked | B*** | S*** | B** | **B** | B*** | S*** | S*** | **S** |
| 8.2 2002, single-banked | tie | tie | tie | **B,S,MH** | B*** | S*** | S*** | **S** |
| 8.3 2002, multi-banked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| 9.1 2000, unbanked | B*** | S*** | tie | **S,B** | B*** | S*** | S*** | **S** |
| 9.2 2000, single-banked | B*** | S* | B*** | **B** | B*** | S*** | B*** | **B** |
| 9.3 2000, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | B** | **B** |
| 9.4 1997, unbanked | B*** | S*** | tie | **B,S** | B*** | S*** | tie | **S,B** |
| 9.5 1997, single-banked | B*** | S*** | tie | **B,S** | B*** | S*** | tie | **B,S** |
| 9.6 1997, multi-banked | tie | tie | tie | **MH,B,S** | B*** | tie | B*** | **B** |
| *V. Large and old firms, continuing banking status; (k,I,q) cross-sections* | | | | | | | | |
| 10.1 2007, unbanked | MH** | MH*** | tie | **MH** | tie | tie | tie | **B,S,MH** |
| 10.2 2007, single-banked | MH** | MH*** | tie | **MH** | MH*** | MH*** | B*** | **MH** |
| 10.3 2007, multi-banked | B*** | S*** | B*** | **B** | B*** | S** | B*** | **B** |
| 10.4 2004, unbanked | MH*** | MH** | S** | **MH** | MH*** | MH*** | B*** | **MH** |
| 10.5 2004, single-banked | MH*** | MH*** | S*** | **MH** | B*** | MH*** | B*** | **B** |
| 10.6 2004, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |
| 11.1 2002, unbanked | MH*** | MH*** | tie | **MH** | B*** | S*** | tie | **S,B** |
| 11.2 2002, single-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 11.3 2002, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |
| 12.1 2000, unbanked | B*** | S* | B*** | **B** | B*** | S*** | S*** | **S** |
| 12.2 2000, single-banked | B*** | S*** | B** | **B** | B*** | S*** | S*** | **S** |
| 12.3 2000, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | tie | **S,B** |
| 12.4 1997, unbanked | B*** | S** | B** | **B** | B*** | S*** | tie | **S,B** |
| 12.5 1997, single-banked | B*** | S*** | B* | **B** | B* | S*** | S*** | **S** |
| 12.6 1997, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |

Notes: The listed regime(s) is best fitting. ***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively.

## Table 5C - Vuong test model comparisons - Industry

| | A. family firms | | | | B. non-family firms | | | |
|---|---|---|---|---|---|---|---|---|
| Comparison: | MH v B | MH v S | B v S | Best fit | MH v B | MH v S | B v S | Best fit |
| *VIII. Construction, continuing banking status; (k,I,q) cross-sections* | | | | | | | | |
| 2007, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | tie | **S,B** |
| 2007, single-banked | B*** | MH*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 2007, multi-banked | B*** | MH*** | B*** | **B** | B*** | MH*** | B*** | **B** |
| 2004, unbanked | B*** | MH** | B*** | **B** | B*** | S*** | B*** | **B** |
| 2004, single-banked | B*** | tie | B* | **B** | B*** | S*** | tie | **S,B** |
| 2004, multi-banked | B*** | MH*** | B*** | **B** | B*** | MH*** | B*** | **B** |
| 2002, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | tie | **S,B** |
| 2002, single-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 2002, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 2000, unbanked | B*** | S*** | tie | **B,S** | MH*** | S*** | S*** | **S** |
| 2000, single-banked | B*** | S*** | B** | **B** | B*** | S*** | S*** | **S** |
| 2000, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 1997, unbanked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| 1997, single-banked | B*** | S* | tie | **B,S** | B*** | S*** | S*** | **S** |
| 1997, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |
| *IX. Non-construction, continuing banking status; (k,I,q) cross-sections* | | | | | | | | |
| 2007, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 2007, single-banked | B*** | MH*** | B*** | **B** | B*** | S*** | B*** | **B** |
| 2007, multi-banked | B*** | S*** | B*** | **B** | B*** | MH*** | B*** | **B** |
| 2004, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 2004, single-banked | MH*** | MH*** | S*** | **MH** | B*** | MH*** | B*** | **B** |
| 2004, multi-banked | MH*** | MH*** | B*** | **MH** | B*** | MH*** | B*** | **B** |
| 2002, unbanked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| 2002, single-banked | MH*** | MH*** | tie | **MH** | B*** | S*** | S*** | **S** |
| 2002, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 2000, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 2000, single-banked | B*** | S*** | tie | **B,S** | B*** | S*** | S*** | **S** |
| 2000, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 1997, unbanked | B*** | S*** | tie | **B,S** | B*** | S*** | S*** | **S** |
| 1997, single-banked | B*** | S*** | B*** | **B** | B*** | S*** | tie | **S,B** |
| 1997, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |

Notes: The listed regime(s) is best fitting. ***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively.

## Table 5D - Vuong test model comparisons - Region

| Comparison: | A. family firms MH v B | MH v S | B v S | Best fit | B. non-family firms MH v B | MH v S | B v S | Best fit |
|---|---|---|---|---|---|---|---|---|
| *VI. Region 1, continuing banking status; (k,I,q) cross-sections* | | | | | | | | |
| 2007, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |
| 2007, single-banked | B*** | MH*** | B*** | **B** | B*** | S*** | B*** | **B** |
| 2007, multi-banked | B*** | S*** | B*** | **B** | B*** | MH*** | B*** | **B** |
| 2004, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |
| 2004, single-banked | B*** | MH*** | B*** | **B** | B*** | S*** | B*** | **B** |
| 2004, multi-banked | tie | MH*** | B*** | **B,MH** | B*** | MH*** | B*** | **B** |
| 2002, unbanked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| 2002, single-banked | MH*** | MH*** | tie | **MH** | B*** | S*** | S*** | **S** |
| 2002, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | tie | **S,B** |
| 2000, unbanked | tie | MH** | B*** | **B,MH** | B*** | S*** | S*** | **S** |
| 2000, single-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 2000, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 1997, unbanked | B*** | S*** | B** | **B** | B*** | S*** | S*** | **S** |
| 1997, single-banked | B*** | S*** | tie | **B,S** | B*** | S*** | B*** | **B** |
| 1997, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |
| *VII. Region 2, continuing banking status; (k,I,q) cross-sections* | | | | | | | | |
| 2007, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 2007, single-banked | B*** | S** | B*** | **B** | B*** | S*** | S*** | **S** |
| 2007, multi-banked | B*** | S*** | B*** | **B** | B*** | MH*** | B*** | **B** |
| 2004, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 2004, single-banked | B*** | MH*** | B*** | **B** | B*** | S*** | B*** | **B** |
| 2004, multi-banked | MH*** | MH*** | B*** | **MH** | MH*** | MH*** | B*** | **MH** |
| 2002, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 2002, single-banked | B*** | tie | B*** | **B** | B*** | S*** | S*** | **S** |
| 2002, multi-banked | B*** | S*** | B*** | **B** | MH*** | S*** | S*** | **S** |
| 2000, unbanked | B*** | S** | tie | **B,S** | B*** | S*** | tie | **B,S** |
| 2000, single-banked | B*** | S*** | B** | **B** | B*** | S*** | S** | **S** |
| 2000, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 1997, unbanked | B** | S*** | tie | **B,S** | B*** | S*** | S*** | **S** |
| 1997, single-banked | B*** | S*** | B*** | **B** | B*** | S*** | B* | **B** |
| 1997, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |

Notes: The listed regime(s) is best fitting. ***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively.

## Table 6 - Vuong Test Model Comparisons - Family Firm Sub-categories

| Comparison: | A. pure family firms MH v B | MH v S | B v S | Best fit | B. family networked firms MH v B | MH v S | B v S | Best fit | C. non-family firms MH v B | MH v S | B v S | Best fit |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *1. 2004-07 data; investment (I) three-year panel* | | | | | | | | | | | | |
| 1.1 04-06, unbanked | B*** | S*** | S*** | **S** | B*** | S*** | tie | **B,S** | MH*** | S*** | S*** | **S** |
| 1.2 04-06, single-banked | MH*** | MH*** | B*** | **MH** | MH* | MH* | tie | **MH** | B*** | S*** | S*** | **S** |
| 1.3 04-06, multi-banked | B*** | S*** | B*** | **B** | tie | tie | tie | **MH,B,S** | MH*** | MH*** | B*** | **MH** |
| *2. 2004-07 data; investment and cashflow (I,q) two-year panels* | | | | | | | | | | | | |
| 2.1 04,05 unbanked | B*** | S*** | S*** | **S** | B*** | S* | B*** | **B** | B*** | S*** | S*** | **S** |
| 2.2 04,05 single-banked | B*** | MH*** | B*** | **B** | MH*** | MH*** | B** | **MH** | tie | S*** | S*** | **S** |
| 2.3 04,05 multi-banked | MH*** | MH*** | B*** | **MH** | tie | tie | tie | **MH,B,S** | MH*** | MH*** | B*** | **MH** |
| 2.4 04,07 unbanked | B*** | S*** | S*** | **S** | B*** | S*** | B** | **B** | B*** | S*** | S*** | **S** |
| 2.5 04,07 single-banked | B*** | S*** | tie | **S,B** | MH*** | MH*** | B** | **MH** | tie | S*** | S*** | **S** |
| 2.6 04,07 multi-banked | MH*** | MH*** | B*** | **MH** | tie | tie | tie | **MH,B,S** | MH*** | MH*** | S*** | **MH** |
| *3. 2004-07 data; assets, investment and cashflow (k,I,q) cross-sections* | | | | | | | | | | | | |
| 3.1 2007, unbanked | B*** | S*** | S*** | **S** | B** | S* | B** | **B** | MH*** | S*** | S*** | **S** |
| 3.2 2007, single-banked | B*** | tie | B*** | **B** | tie | MH* | tie | **MH,B** | B*** | S*** | tie | **S,B** |
| 3.3 2007, multi-banked | B*** | S*** | B*** | **B** | tie | tie | tie | **MH,B,S** | B*** | MH*** | B*** | **B** |
| 3.4 2004, unbanked | B*** | S*** | tie | **S,B** | B*** | S*** | B** | **B** | B*** | S*** | B*** | **B** |
| 3.5 2004, single-banked | B*** | S*** | B*** | **B** | MH*** | MH*** | tie | **MH** | B*** | S*** | B*** | **B** |
| 3.6 2004, multi-banked | B*** | MH*** | B*** | **B** | B* | tie | B* | **B** | B*** | MH*** | B*** | **B** |
| *4. 2001-03 data; (k,I,q) cross-sections* | | | | | | | | | | | | |
| 4.1 2002, unbanked | B*** | S*** | S*** | **S** | MH*** | MH** | tie | **MH** | B*** | S*** | S*** | **S** |
| 4.2 2002, single-banked | B*** | tie | B*** | **B** | tie | tie | tie | **MH,B,S** | B*** | S*** | S*** | **S** |
| 4.3 2002, multi-banked | B*** | S*** | B*** | **B** | tie | tie | tie | **MH,B,S** | B** | S*** | S*** | **S** |
| *5. 1997-00 data; (k,I,q) cross-sections* | | | | | | | | | | | | |
| 5.1 1997, unbanked | B*** | S*** | S*** | **S** | B** | S*** | tie | **B,S** | MH** | S*** | S*** | **S** |
| 5.2 1997, single-banked | B*** | S*** | B*** | **B** | B*** | S*** | tie | **B,S** | B** | S*** | S** | **S** |
| 5.3 1997, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |
| *6. (k,I,q) cross-sections, any continuing banking status* | | | | | | | | | | | | |
| 6.1 2004, small and young | B*** | S*** | S*** | **S** | B*** | S*** | B** | **B** | B*** | S*** | S*** | **S** |
| 6.2 2004, large and old | B*** | S*** | B*** | **B** | tie | tie | tie | **MH,B,S** | B*** | S*** | B*** | **B** |

The listed regime(s) is best fitting. ***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively.

## Table 7 - Vuong Test Model Comparisons - Robustness

| Comparison: | A. family firms MH v B | MH v S | B v S | Best fit | B. non-family firms MH v B | MH v S | B v S | Best fit |
|---|---|---|---|---|---|---|---|---|
| *1. Continuing banking status, no balanced panel* | | | | | | | | |
| 1.1 2004, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 1.2 2004, single-banked | MH*** | MH*** | S*** | **MH** | MH*** | S*** | S*** | **S** |
| 1.3 2004, multi-banked | B*** | S*** | tie | **B,S** | B*** | B*** | B*** | **B** |
| 1.4 2000, unbanked | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| 1.5 2000, single-banked | MH*** | MH*** | S*** | **MH** | B*** | S*** | S*** | **S** |
| 1.6 2000, multi-banked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| *2. Switching bank status* | | | | | | | | |
| 2.1 2004 all firms | tie | MH*** | B*** | **MH,B** | MH*** | MH*** | B*** | **MH** |
| 2.2 2007 all firms | MH*** | MH*** | B*** | **MH** | B*** | MH*** | B*** | **B** |
| *3. Any continuing banking status* | | | | | | | | |
| 3.1 2004, all firms | MH*** | MH*** | B*** | **MH** | B*** | MH*** | B*** | **B** |
| 3.2 2004, small and young | B*** | S*** | S*** | **S** | B*** | S*** | S*** | **S** |
| 3.3 2004, large and old | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |
| *4. Placebo family status* | | | | | | | | |
| 4.1 2004, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | B*** | **B** |
| 4.2 2004, single-banked | B*** | S*** | B*** | **B** | B*** | MH*** | B*** | **B** |
| 4.3 2004, multi-banked | MH*** | MH*** | B*** | **MH** | MH*** | MH*** | B*** | **MH** |
| *5. Risk aversion* | | | | | | | | |
| 5.1 2004, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | tie | **B,S** |
| 5.2 2004, single-banked | B*** | S*** | tie | **B,S** | B*** | S*** | S*** | **S** |
| 5.3 2004, multi-banked | MH*** | MH*** | tie | **MH** | MH*** | MH*** | tie | **MH** |
| *6. Adjustment costs* | | | | | | | | |
| 6.1 2004, unbanked | B*** | S*** | tie | **B,S** | B*** | S*** | S*** | **S** |
| 6.2 2004, single-banked | tie | MH*** | B*** | **B,MH** | B*** | S*** | tie | **B,S** |
| 6.3 2004, multi-banked | B*** | S*** | tie | **B,S** | B*** | S*** | tie | **B,S** |
| *7. (k',I,q) cross-section data from 2004-05* | | | | | | | | |
| 7.1 2004, unbanked | B*** | S*** | B*** | **B** | B*** | S*** | S*** | **S** |
| 7.2 2004, single-banked | MH*** | MH*** | B*** | **MH** | B*** | MH*** | B*** | **B** |
| 7.3 2004, multi-banked | tie | MH*** | B*** | **MH,B** | B*** | MH*** | B*** | **B** |

Notes: All specifications use (k,I,q) cross-section data from the listed year, unless specified otherwise. The listed regime(s) is best fitting. ***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively.

## Appendix Table A1 - More sample comparisons

| | (1) | (2) | (3) |
|---|---|---|---|
| **1997-2000** | | | |
| observations | 841951 | 353361 | 216060 |
| number of firms | 328336 | 137591 | 54015 |
| mean age | 10.6 | 11.7 | 13.7 |
| mean assets, k | 906 | 1444 | 1759 |
| median k | 86 | 129 | 201 |
| stdev k | 21340 | 31194 | 21490 |
| mean cash flow, q | 300 | 506 | 678 |
| median q | 43 | 65 | 106 |
| stdev q | 6882 | 10424 | 8372 |
| mean investment, I | 194 | 296 | 388 |
| median I | 13 | 20 | 34 |
| stdev I | 6335 | 6759 | 5571 |
| percent family firms | 35% | 37% | 45% |
| percent unbanked | 19% | 17% | 9% |
| percent single-banked | 28% | 11% | 7% |
| percent multi-banked | 53% | 72% | 84% |
| | | | |
| **2001-2003** | | | |
| observations | 1093904 | 427835 | 261687 |
| number of firms | 512052 | 211619 | 87229 |
| mean age | 10.5 | 11.4 | 13.7 |
| mean assets, k | 927 | 1433 | 2051 |
| median k | 73 | 100 | 183 |
| stdev k | 30064 | 45876 | 56698 |
| mean cash flow, q | 276 | 482 | 721 |
| median q | 32 | 42 | 83 |
| stdev q | 9683 | 15319 | 18658 |
| mean investment, I | 182 | 256 | 369 |
| median I | 9 | 11 | 24 |
| stdev I | 6587 | 7654 | 8927 |
| percent family firms | 31% | 33% | 42% |
| percent unbanked | 23% | 26% | 15% |
| percent single-banked | 31% | 16% | 11% |
| percent multi-banked | 46% | 58% | 74% |

Notes:

(1) all observations with non-missing (k,I,q) data from any of the listed years

(2) observations of firms with non-missing (k,I,q) data from any of the listed years and continuing bank status

(3) observations of firms with non-missing (k,I,q) data present in all listed years and continuing bank status (balanced panel)

* All means, medians, standard deviations (stdev) and percentages are computed at the observation level.